

Indicatoren voor risicodetectie in het Nederlandse onderwijs

Kwaliteitszorg, financiële ratio's en verticale integratie



Jos Verkroost

***Indicatoren voor risicodetectie in het Nederlandse
onderwijs
Kwaliteitszorg, financiële ratio's en verticale
integratie***

***Indicators for risk detection in Dutch Education
Quality management, financial ratios and vertical
integration***

Proefschrift

**ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus**

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

**De openbare verdediging zal plaatsvinden op
vrijdag 1 juli 2016 om 13.30 uur**

**Joseph Jacobus Hendrikus Verkroost
geboren te Someren**

Promotiecommissie

Promotor:

Prof.dr. G.W.J. Hendrikse

Overige leden:

Prof.dr. S.P. Kaptein

Prof.dr. A. de Jong

Prof.dr. I.F. de Wolf

Voor Annemiek

Voorwoord

Dit onderzoek is gestart in het kader van een initiatief van Muel Kaptein en Martin de Bree om wetenschappelijk onderzoek te stimuleren naar toezicht. Het is een groot voordeel gebleken om als buitenpromovendus in een dergelijk groepsverband te opereren. Het onderlinge contact met de groepsgenoten, de reflectie op elkaars tussenproducten alsmede de mogelijkheid om zaken georganiseerd te krijgen, hebben zonder meer bijgedragen aan de voortgang van het proefschrift dat nu voorligt.

De spil in de totstandkoming daarvan is evenwel George Hendrikse, mijn promotor, geweest. Waar ik in mijn oorspronkelijke promotievoorstel al wel ideeën had om een verbinding te leggen tussen toezicht op het onderwijs en bedrijfskundige en bedrijfseconomische principes, was het zijn inbreng die geleid heeft tot een benadering en verdieping die ik vooraf niet had kunnen vermoeden. Daarbij was hij zeer stimulerend in zijn kritische, inhoudelijke reflectie op de tussenproducten die ik voorlegde. Dit heeft in hoge mate bijgedragen aan een voorspoedige voortgang van het werk. Daarvoor mijn grote dank.

Waar George bovendien zelf van mening was aan de grenzen van zijn eigen deskundigheid te zitten, heeft hij actief collega-deskundigen ingeschakeld om vanuit hun specialistische invalshoek hun licht op onderdelen te laten schijnen. Zeer waardevol commentaar is daarbij geleverd door Bart Dietz op het vlak van de statistische verwerking en Philip Fliers op het terrein van de financiële ratioanalyse.

Verder ben ik dank verschuldigd aan de Inspectie van het Onderwijs en een aantal collega's daar. De inspectie heeft mij in staat gesteld mijn werkzaamheden te verrijken met het onderzoek dat ik heb kunnen verrichten. Bovendien was ik in de gelegenheid gebruik te maken van een aantal van de noodzakelijke gegevens daarvoor. Meerdere collega's, in het bijzonder van de directies Rekenschap en Kennis, hebben op verschillende momenten op delen van het onderzoek gereageerd en daarmee bijgedragen aan de inhoud ervan.

Mijn dank geldt evenzeer de commissieleden die bereid zijn geweest mijn proefschrift te beoordelen en van aanvullend en verdiepend commentaar te voorzien.

Tenslotte wil ik mijn echtgenote, Annemiek, aan wie dit proefschrift is opgedragen, bedanken. Hoewel ik mij voor mijn werk nooit afzonder en dus aanspreekbaar ben gebleven, legt een dergelijk traject toch een flink beslag op aandacht en tijd die er anders zou zijn voor andere activiteiten. Dat het voorspoedig is verlopen, is daarom in belangrijke mate te danken aan haar bereidheid wat in te schikken en gelijktijdig een stimulerende belangstelling aan de dag te leggen.

Inhoud

Voorwoord	5
1 Inleiding	11
1.1 Principaal-agentrelaties.....	11
1.2 De overheid en de uitvoeringsorganisaties in het onderwijs	13
1.3 Toezicht	16
1.4 Risicogericht toezicht	19
1.5 Risicogericht toezicht bij de onderwijsinspectie ..	21
1.6 Effecten en neveneffecten van sturing en toezicht.....	22
1.7 Te onderzoeken vraagstukken	25
Literatuur	27
2 Voorspelt kwaliteitszorg de prestaties bij instellingen voor voortgezet onderwijs?	31
Samenvatting	31
2.1 Inleiding	31
2.2 Literatuuronderzoek	33
2.2.1 Total Quality Management	33
2.2.2 Kwaliteitsmanagement en prestaties.....	35
2.2.3 Kwaliteitsmanagement in het onderwijs	38
2.2.4 Wettelijke eisen aan kwaliteitszorg in het voortgezet onderwijs	39
2.2.5 TQM versus de werkwijze van de inspectie	41
2.2.6 Kwaliteitszorg en prestaties in het onderwijs in Nederland	44
2.3 Vraagstelling en hypothese.....	45
2.4 De onderzoeksgegevens.....	48
2.4.1 Instellingen, financiën en onderwijsopbrengsten	48
2.4.2 De kwaliteitszorgdata.....	50
2.4.3 Bewerking van de gegevens	51
2.5 Onderzoeksbevindingen	55

2.5.1	Voorspelling van het financieel presteren.....	55
2.5.2	Voorspelling van de opbrengsten.....	56
2.5.3	Aanvullende analyses op schoolniveau	58
2.5.4	Voorspelling van de prestaties in het algemeen.....	59
2.6	Conclusie en discussie	60
	Literatuur	62
3	Voorspellen van financieel disfunctioneren van scholen.....	65
	Samenvatting	65
3.1	Inleiding	65
3.1.1	Aanleiding	65
3.1.2	Vraagstelling.....	67
3.1.3	Opbouw van het hoofdstuk	68
3.2	Theoretische overwegingen	69
3.2.1	Fouten van type 1 en type 2	69
3.2.2	Keuze van de indicatoren	71
3.2.3	Analysemethoden.....	73
3.2.4	Besluitvorming volgens verschillende processen	75
3.2.5	Modellen voor de voorspelling van financiële kwetsbaarheid	78
3.2.6	Modellen voor profitorganisaties	78
3.2.7	Een model voor non-profitorganisaties	80
3.2.8	Het toezichtmodel van de onderwijsinspectie .	82
3.3	Onderzoeksgegevens	85
3.3.1	Beschrijvende statistiek	86
3.3.2	Wijze van analyseren van de gegevens	88
3.4	Onderzoeksbevindingen	90
3.4.1	Resultaten van de regressies	90
3.4.2	Voorspelling van de resultaten op basis van jaarrekening 2012.....	92
3.4.3	Aanpassing van de onderscheidingswaarde....	94

3.4.4	Gecombineerde besluitvormingsmethoden	95
3.4.5	Resultaten van de verschillende methoden	96
3.4.6	Analyse van de type-2-fouten	99
3.5	Conclusie en discussie	101
	Literatuur	103
4	Waardecreatie door verticale integratie tussen po, vo en mbo.	107
	Samenvatting	107
4.1	Inleiding	107
4.2	Verticale integratie	108
4.3	Verticale integratie in het onderwijs in Nederland	115
4.4	Onderwijsintegratie in de economische theorie..	121
4.5	Vraagstelling	124
4.6	Kwantitatieve analyse	125
4.6.1	Onderzoeksdata	125
4.6.2	Beschrijvende statistiek	126
4.6.3	Wijze van analyseren	127
4.6.4	Doorstroom van leerlingen	128
4.6.5	Voortijdig schoolverlaten	133
4.6.6	Financieel presteren.....	134
4.7	Conclusie en discussie	136
	Literatuur	137
5	Conclusies en discussie.....	141
5.1	Beoordeling van de kwaliteitszorg	141
5.2	Financieel disfunctioneren	143
5.3	Verticale integratie	146
5.4	Neveneffecten van de indicatoren.....	147
5.5	Toepassingen in de praktijk	148
5.6	Algemeen.....	150
	Literatuur	151
	Samenvatting	155

Abstract	157
Curriculum vitae.....	159
Portfolio	161

1 Inleiding

De positie van een overheidsinspectie met toezicht op publieke dienstverlening, zoals de onderwijsinspectie, is er een in de relatie tussen opdrachtgevers en uitvoerders. De opdrachtgever is in deze situatie het ministerie van Onderwijs, Cultuur en Wetenschappen (OCW) en de uitvoerders zijn de schoolbesturen. Om in de gaten te houden of de schoolbesturen de taak die zij hebben, tot tevredenheid uitvoeren, ziet de onderwijsinspectie op de scholen toe.

Opdrachtgevers en uitvoerders worden in de literatuur ook wel aangeduid als principalen en agenten (Gibbons, 2016a). Waar de principaal niet zelf in staat is de werkzaamheden van de agenten te bewaken, zijn mechanismen nodig om grotere zekerheid te verschaffen op de gewenste uitvoering van de werkzaamheden. Dit leidt tot agency-kosten (Shapiro, 2005) of verificatiekosten (Dixit, 2002). Deze kosten zijn een verzamelnaam voor alles wat er in een relatie mis kan gaan. Om deze kosten te verminderen, worden voorzieningen getroffen. Inspectiediensten in het algemeen, en binnen de context van het toezicht op het onderwijs de onderwijsinspectie, zijn als zo'n voorziening te beschouwen. Zij toetsen aan de hand van een aantal richtlijnen of de agenten de bedoelde publieke dienstverlening daadwerkelijk leveren. In het geval van het onderwijs is dat onderwijs van voldoende kwaliteit.

Om dat toezicht effectief en efficiënt te kunnen houden, is toetsing van de resultaten nodig aan beoogde doelen en zijn indicatoren nodig om te bepalen welke agenten het beter of juist minder goed doen, teneinde gericht te kunnen interveniëren. Dergelijke indicatoren moeten zoveel mogelijk zeggen over (risico's voor) het bereiken van de gestelde doelen door, maar gelijktijdig zo min mogelijk neveneffecten veroorzaken in het handelen van de agent.

1.1 Principaal-agentrelaties

De algemene gedachtegang achter het principaal-agent model is dat er sprake is van een persoon of instantie die een bepaald resultaat wil bereiken en niet in staat is dat (volledig) zelf te doen (Gibbons, 2016a). Daarvoor is het nodig een uitvoerder aan te zoeken die dat werk doet. Op dat moment is er sprake van een verhouding van principaal en agent. De principaal geeft de agent opdrachten, de agent voert die uit en wordt, afhankelijk van de vervulling van de opdrachten, beloond. De werkelijkheid is aanzienlijk complexer dan dit model suggereert, maar de essentie ervan is wel geschikt om deze nader te beschouwen.

Een eerste complicatie in deze relatie is dat de belangen van de principaal en agent niet (volledig) samenvallen. Om ervoor te zorgen dat de

belangen worden gestroomlijnd, biedt de principaal een contract aan dat dit dient te bewerkstelligen. Afhankelijk van de omstandigheden kunnen dat volledige en onvolledige contracten zijn. Volledige contracten veronderstellen namelijk dat alle activiteiten en resultaten van de agent en de handelingen van de principaal vooraf zodanig te definiëren zijn dat een derde partij (veelal de rechter) in staat is te beoordelen of al dan niet aan de contractvoorwaarden is voldaan. De praktijk maakt het vaak onmogelijk dergelijke contracten te sluiten in verband met onzekerheden rond de uitvoering of complexiteit van de uit te voeren opdrachten. In dat geval moet worden volstaan met onvolledige contracten (Hendrikse, 2003). Daarin zijn niet alle contractvoorwaarden vastgelegd, maar is er sprake van veel globalere afspraken die niet door een derde partij getoetst kunnen worden (Baker, Gibbons en Murphy 2002). Dergelijke contracten vergen vertrouwen tussen de contractpartijen en functioneren vaak in langduriger relaties. Gibbons (2016b) spreekt in dat verband van relationele contracten.

Een tweede complicatie is dat er bij het sluiten en uitvoeren van de contracten sprake is van informatieasymmetrie tussen de principaal en de agent. Dixit (2002) onderscheidt drie typen van deze informatieasymmetrie. Deze typen sluiten elkaar niet uit, maar kunnen ook gelijktijdig op een situatie van toepassing zijn. Het eerste is moral hazard. Moral hazard doet zich voor gedurende de relatie tussen de principaal en de agent, dus nadat een contract is gesloten. De principaal kan niet precies zien wat de agent doet, maar kan alleen de resultaten van zijn handelen waarnemen. Dixit maakt daarbij nog een onderscheid tussen resultaten die alleen observeerbaar zijn en resultaten die ook verifieerbaar zijn. In het laatste geval kan ook een derde partij over het resultaat oordelen, hetgeen een volledig contract veronderstelt.

Het tweede type is dat van de adverse selection. Deze vorm van asymmetrie is aan de orde voor aanvang van het contract tussen de principaal en de agent. Een agent met voorkennis die de principaal niet heeft, is in een gunstiger positie bij het afsluiten ervan. Zo kan hij de omstandigheden ongunstiger voorstellen, waardoor hij betere voorwaarden kan bedingen. Tenslotte noemt Dixit costly verification. Hierbij is het resultaat van het handelen van de agent alleen tegen hoge kosten te bepalen. De principaal staat dan voor de keuze deze kosten te maken, dan wel de agent op zijn woord te geloven.

Een belangrijke verbreding wordt beschreven door Dixit (2002) en ook Shapiro (2005). Dat is het gegeven dat er eigenlijk nooit sprake is van een contractsituatie tussen precies een principaal en een agent die zich verder geïsoleerd van de omgeving afspeelt. Volgens Dixit is dat in publieke organisaties nog in sterkere mate het geval dan in de private sector. Agenten hebben te maken met meerdere principalen en met uiteenlopende doelen die onderling met elkaar kunnen concurreren. Soms kunnen de rollen van principaal en agent wisselen bij verschillende situaties waarin men met elkaar opereert. Bovendien is er vaak sprake

van een gelaagde situatie, waarin meer niveaus een keten van principalen en agenten vormen. Dixit (2002) typeert onder meer de situatie waarin een toezichthouder namens een finale principaal optreedt, als zo'n gelaagde situatie. Tirole (1986) laat zien dat zich daarbij situaties voor kunnen doen als samenspanning tussen toezichthouder en agent en tekortschietende doorzettingsmacht voor de toezichthouder in zijn tussenpositie.

Een tweede belangrijke verbreding van het principaal-agent model is dat de nauwkeurigheid waarmee prestaties worden gemeten, gevolgen heeft voor de omvang van de prestatiebeloning voor de agent (Baker, 2000). De intensiteit van de prestatieprikkels voor de agent zal sterker zijn naarmate de wijze waarop de prestatieprikkels worden gemeten meer overeenkomt met de doelstellingen van de organisatie. Zo hecht de overheid sterk aan de verbetering van het niveau van taal en rekenen als basis voor de verdere ontwikkeling van leerlingen, en heeft zij deze daarom ook vastgelegd als prestatie maatstaf. Het blijkt dat die intensiteit ook door de agenten (scholen) zo wordt ervaren (zie ook paragraaf 1.6).

1.2 De overheid en de uitvoeringsorganisaties in het onderwijs

"Het onderwijs is voorwerp van aanhoudende zorg der regering" (Art. 23 lid 1 Grondwet). Met dit artikel positioneert de overheid zich als principaal van het Nederlandse onderwijs. Zij stelt vervolgens dat dit onderwijs vrij is, behoudens het toezicht door de overheid (Art. 23 lid 2 Grondwet). De overheid draagt zorg voor openbaar onderwijs en bekostigt bijzonder onderwijs naar dezelfde maatstaven. Met dit uitgangspunt neemt de overheid gelijktijdig de zorg op zich voor het functioneren van een groot aantal agenten die belast zijn, dan wel zichzelf belasten met het uitvoeren van onderwijs dat aan een aantal bij wet te regelen deugdelijkheidsmaatstaven voldoet.

Het contract dat het ministerie op basis hiervan met de schoolbesturen sluit, is in termen van de principaal-agentrelatie er een van onvolledige contractering. Dat wordt formeel al gesteld door de bepaling dat het onderwijs vrij is in lid 2. Het is echter ook in praktische zin onmogelijk het totaal van de maatschappelijke opbrengsten waaraan een bestuur een bijdrage geacht wordt te leveren, neer te leggen in een volledig contract. Of leerlingen goede burgers worden die een productieve bijdrage leveren aan de samenleving, en wat daaraan de precieze bijdrage is van een school of schoolbestuur, is niet geheel in afspraken vast te leggen die rechtens afdwingbaar zijn, en dat zou ook in strijd zijn met het hiervoor aangehaalde artikel 23 uit de Grondwet. Dat betekent dat er een grote

ruimte in de contracten zit en dat een belangrijk deel van de prestaties is gebaseerd op wederzijds vertrouwen.

De basis voor dit vertrouwen wordt gedeeltelijk gevormd door vastlegging van een aantal relevant geachte randvoorwaarden voor de uitvoering van de hoofdtaak. In de grondwet is de voorwaarde voor het stellen van dergelijke eisen geregeld, waar wordt gesteld dat bij wet eisen van deugdelijkheid kunnen worden bepaald voor onderwijs dat geheel of gedeeltelijk uit de openbare kas wordt bekostigd (Artikel 23, lid 5 Grondwet). In de praktijk zijn deze eisen vastgelegd in zogenaamde sectorwetten, zoals de wet op het primair onderwijs (WPO) (OCW, 1981). Het is daarbij overigens een kwestie van perspectief of er sprake is van een complex maar onvolledig contract, dan wel van een serie contracten met soms een volledig en anders een onvolledig karakter. De wettelijke voorschriften zijn immers in rechte afdwingbaar.

Ook al is er altijd sprake geweest van onvolledige contractering, toch is de verhouding tussen het aantal formele en informele afspraken in de loop van de tijd aan veranderingen onderhevig geweest. Nog tot ver in de jaren tachtig kende de onderwijsregelgeving een zeer groot aantal, zeer gedetailleerde, wettelijke voorschriften waaraan scholen zich te houden hadden. Weliswaar was het contract, als het als een geheel wordt beschouwd, in totaal onvolledig, maar de beperkingen in met name de voorwaardelijke sfeer waren groot. Een kernpunt daarin was het systeem van de zogenaamde declaratiebekostiging die in die tijd gemeengoed was in het onderwijs (Knippenberg en Van der Ham, 1994). Dit systeem houdt in dat schoolbesturen binnen een uitgebreid stelsel van normen alle kosten die ze maken, precies vergoed krijgen, van inventaris tot personeel. Verwerving en besteding zijn in een dergelijk systeem precies voorgeschreven. De gedachtegang hierbij was dat scholen binnen die context van randvoorwaardelijke waarborgen bijna automatisch onderwijs van goede kwaliteit zouden leveren.

In essentie leidde deze verhouding van afdwingbare en niet afdwingbare voorschriften in de contractering ertoe dat schoolbesturen en scholen vooral de organisatorisch context schiepen voor het aanbieden van onderwijs, maar feitelijk nauwelijks verantwoordelijk konden worden gesteld voor het inhoudelijk resultaat daarvan. Aan de veranderende eisen aan het onderwijs vanuit de samenleving kon, binnen de strakke omgeving van de wettelijke eisen, onvoldoende tegemoet worden gekomen (Elte en Scholtes, 2001). Dat is aanleiding geweest tot een aantal veranderingen die vanaf de jaren tachtig in gang werden gezet, en door minister Deetman als volgt werden getypeerd:

“Het terugtreden van de overheid in meer globale regelgeving en in het aansporen van eigen meesterschap gaat dus vergezeld van een nadrukkelijk toetreden op punten als kwaliteitsbewaking, eindtermen, externe democratisering en het evenwicht van de drie wezenlijke functies

van het onderwijs, namelijk individuele ontplooiing, cultuuroverdracht en voorbereiding op de beroepsuitoefening.¹"

In de kern werd het systeem van declaratiebekostiging geleidelijk vervangen door zogenaamde lumpsum-bekostiging. In dat systeem ontvangen schoolbesturen op basis van een aantal normen, waaronder leerlingenaantal, een bedrag dat zij vervolgens binnen grenzen vrij kunnen besteden. Verwerving van de middelen ligt nog altijd vast, maar de besteding ervan is in hoge mate vrij. Met de verandering van de bekostiging is daarom een groot aantal andere voorschriften geschrapt of aangepast om die vrijheid ook in de praktijk mogelijk te maken.

Deze ontwikkelingen leidden tot een ingrijpende aanpassing van de contracten. Deze bleven onvolledig, maar de inhoud ervan werd geheel anders. Scholen en besturen werden in toenemende mate zelf verantwoordelijk voor de inhoud en de resultaten van het onderwijs en er ontstond een zodanig groot aantal vrijheidsgraden in de inrichting daarvan, dat kwaliteit van het onderwijs als een aan de school toe te rekenen eigenschap kon worden beschouwd en beoordeeld. Het aantal formeel contracteerbare afspraken met de overheid nam dienovereenkomstig af, zoals hiervoor beschreven.

Daarvoor in de plaats kwam evenwel een aantal nieuwe overeenkomsten waardoor het aantal principalen voor onderwijsinstellingen toenam. Dit gold onder meer de arbeidsrechtelijke bepalingen voor het personeel, die voordien centraal door de overheid werden bepaald, maar nadien moesten worden vastgesteld in overleg tussen vertegenwoordigers van de schoolbesturen en de personeelsvakorganisaties, binnen de door de overheid gestelde, budgettaire kaders. In termen van Dixit (2002) is daarbij sprake van een toename van het aantal dimensies in de verhouding tussen principaal en agent. Er zijn meerdere opdrachtgevers met uiteenlopende doelen zodat er in de werkelijkheid sprake is van meerdere principalen. Dit past ook in het beleid, waar gesteld wordt dat onderwijsinstellingen rekening dienen te houden met verschillende groepen belanghebbenden in een horizontale dialoog.

Dit roept de vraag op hoe agenten omgaan met de concurrerende opdrachten die zij van die verschillende principalen krijgen. Buiten die opdrachten kunnen agenten ook te maken krijgen met onderlinge concurrentie en met verschillen in opvattingen over de opdrachten die zij van de principaal ontvangen, zeker waar er sprake is van onvolledige contractering. Het hoeft geen betoog dat van een dergelijke, complexe situatie ook in het Nederlandse onderwijs sprake is.

¹ HSG TK 1987-1988, Rijksbegroting 1988 viii, 16 december 1987, 35, 1883. In: Knippenberg e.a. 1994.

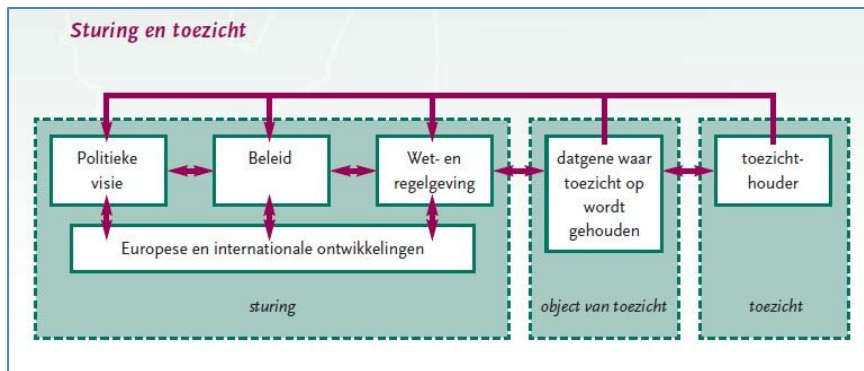
1.3 Toezicht

Onder toezicht wordt in de Kaderstellende visie op toezicht verstaan het verzamelen van informatie over de vraag of een handeling of zaak voldoet aan de daaraan gestelde eisen, het zich vormen van een oordeel daarover en het eventueel naar aanleiding daarvan interveniëren (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK), 2005). Het toezicht kan gericht zijn op drie activiteiten die elkaar kunnen overlappen. Allereerst is dat het naleven van regels. Een tweede is het verschaffen van informatie over de prijs en kwaliteit van (publieke) dienstverlening door zelfstandige organisaties. Tenslotte is dat het informeren van minister, parlement en samenleving over ontwikkelingen in de praktijk en de effecten van beleid (BZK, 2005).

De Kaderstellende visie maakt een onderscheid tussen toezicht en handhaving. Handhaving zou te beschouwen kunnen zijn als een vorm van interventie volgens de definitie van toezicht, zoals gebeurt in Beraad van Inspecteurs-Generaal (2006). De Kaderstellende visie legt deze verbinding niet expliciet. Zij stelt wel dat, waar het gaat om het reguleren van gedrag dat niet streng verboden wordt, nalevingstoezicht en handhaving samenvallen. Velders en Brunia (2013) stellen vast dat beide begrippen nogal wisselend door toezichthouders vanuit traditionele opvattingen worden gebruikt, soms in een vrij smalle betekenis en elders bijna als elkaars synoniemen.

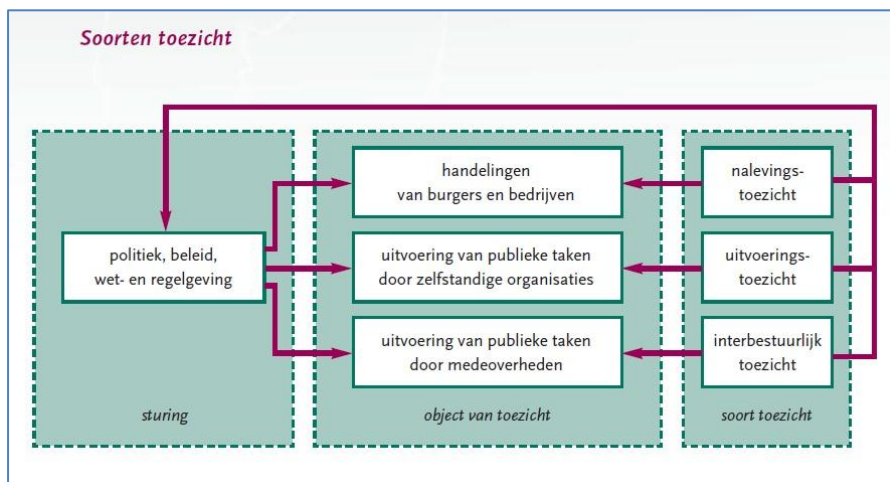
De Kaderstellende visie (BZK, 2005) laat in Figuur 1 de onderlinge relatie zien tussen de wetgever/het ministerie (de principaal, weergegeven door het blokje "sturing"), de toezichthouder en het object van toezicht (de agent, in dit geval de school). De toezichthouder bewaakt namens de overheid als principaal het handelen van de agent. Voor wat betreft de school als het object van toezicht geldt dat het niet strikt gaat om de agent als persoon of organisatie, maar om de agent en zijn handelen voor zover vallend onder het desbetreffende toezicht². De Kaderstellende visie omschrijft het als "datgene waar toezicht op wordt gehouden". De Leidraad objectgericht risicomanagement (Beraad van Inspecteurs-Generaal, 2006) beschrijft objecten van toezicht als de onderwerpen waarop het proces wordt toegepast en geeft daarbij als voorbeelden onder meer bedrijven, bestuursorganen, burgers, milieu en gebouwen. Sommige daarvan zijn als agent aan te merken, maar andere zijn eerder te beschouwen als thema's.

² Zo heeft de onderwijsinspectie dus toezicht op de schoolbesturen voor zover het gaat om hun handelen inzake het onderwijs en de financiën, maar niet op andere zaken die buiten de wettelijke competentie van de inspectie vallen.



Figuur 1 Relatie sturing en toezicht (uit: BZK, 2005)

Relevant is daarnaast het onderscheid naar verschillende soorten toezicht (BZK, 2005). Figuur 2 toont een onderverdeling naar drie categorieën objecten van toezicht en daarmee samenhangend drie soorten van toezicht. De eerste categorie, nalevingstoezicht, geldt voor alle burgers en organisaties, ook voor die lager in het schema. Hierbij gaat het er primair om erop toe te zien dat deze zich aan de wet houden. De daarbij behorende interventie wordt vaker aangeduid als handhaving. Het uitvoeringstoezicht is gericht op de primaire activiteiten van organisaties met een publieke taak. De taken en bevoegdheden van de toezichthouder zijn in het algemeen ook hierop toegesneden. In het kader van de interventie is ook hier handhaving mogelijk, waar de uitvoering van de publieke taak met wettelijke voorschriften is omgeven. Het toezicht op medeoverheden is een laatste categorie. Deze vorm wordt onderscheiden omdat zij niet ondergeschikt zijn aan de centrale overheid en een eigen democratische controle kennen.



Figuur 2 Soorten toezicht (uit: BZK, 2005)

De principaal-agentverhouding is alleen op de tweede categorie van toepassing. De verhouding tussen medeoverheden is, zoals gezegd, in het algemeen niet van een dergelijke aard, met uitzondering van een beperkt aantal specifieke bepalingen. De overige burgers en bedrijven voeren geen publieke taken uit en staan op grond daarvan niet in een dergelijke verhouding tot de overheid.

Met artikel 23, lid 2 van de grondwet wordt voor het onderwijs formeel bepaald dat de overheid toezicht houdt op de uitvoering ervan. Daarmee wordt vastgelegd dat er een intermediaire instantie is in de principaal-agentrelatie tussen het ministerie van OCW en de schoolbesturen, de toezichthouder. De rol van deze toezichthouder is tot 2002 geregeld via bepalingen in de verschillende onderwijswetten, en sinds 2002 in de Wet op het Onderwijstoezicht (OCW, 2002).

Deze onderverdeling in meerdere hiërarchische lagen is in de praktijk vaak aan de orde. Dixit (2002) stelt dat dit in het bijzonder van belang is in de publieke sector en het toezicht is daarvan een typisch voorbeeld. De toezichthouder opereert dan als agent ten opzichte van de wetgever/het ministerie en weer als principaal ten opzichte van de onder toezicht staanden (Tirole, 1986). De Kaderstellende visie (BZK, 2005) maakt dit onderscheid niet, zoals hiervoor te zien is, maar beschouwt de toezichthouder als een verlengstuk van de principaal.

Uitgaande van Dixit (2002) ontstaat er aldus een extra dynamiek in het stelsel waarbij de toezichthouder zich als een eigenstandige partij kan ontpoppen. Daarnaast wordt de publieke sector, aldus Dixit, in hogere mate dan de private sector gekenmerkt door het opereren van meerdere principalen naast elkaar. Het gevolg daarvan is dat de agenten in deze sector ook belast worden met meer taken die leiden tot concurrerende eisen en prioriteiten. Deze situatie wordt in het Nederlandse onderwijs actief gestimuleerd doordat instellingen geacht worden rekening te houden met hun (lokale) belanghebbenden en daaraan ook verantwoording af te leggen. In veel gevallen is dat zelfs wettelijk voorgeschreven en houdt de toezichthouder ook daar toezicht op.

Het toezicht op het onderwijs moet daarmee opereren in een veld waarin sprake is van onvolledige contracten tussen de finale principaal (het ministerie) en de finale agent (de onderwijsinstellingen) en waarin die contractvoorwaarden, ook al zijn het wettelijke eisen, moeten concurreren met andere eisen uit de omgeving. De toezichthouder moet daarin een optimaal toezicht uitoefenen.

1.4 Risicogericht toezicht

Risicogericht³ toezicht is in toenemende mate in de belangstelling gekomen vanaf de eeuwwisseling. Black (2010) omschrijft dit toezicht als de ontwikkeling en het gebruik van systematische, beslissingsondersteunende kaders en procedures om toezichtactiviteiten te prioriteren en toezichtcapaciteit toe te wijzen, gebaseerd op een inschatting van de risico's die organisaties aan de dag leggen in relatie tot de maatschappelijke doelen van de regelgeving.

Black (2010) onderscheidt vier gemeenschappelijke elementen in alle systemen voor risicogericht toezicht. Het eerste is dat het beginpunt om capaciteit toe te wijzen gebaseerd is op risico's en niet op wettelijke voorschriften. De risicogerichtheid moet ertoe bijdragen een verantwoorde en expliciete afweging te maken in de keuze van te handhaven voorschriften uit de schier onuitputtelijke hoeveelheid regelgeving die er bestaat. Het tweede is dat risicogericht toezicht een taxatie maakt van de ernst van de ongewenste gebeurtenissen en de kans dat deze zich voordoen. Op basis daarvan maken de toezichthouders ten derde een score of rangorde van de organisaties of activiteiten waar zij op toezien en delen zij tenslotte toezichtcapaciteit in.

Risicogericht toezicht kan daarmee beschouwd worden als een middel tot reductie van de verificatiekosten (Dixit, 2002). Dixit beschrijft als een van de mogelijkheden daartoe een methode van selectie op basis van een aantal indicatoren van agenten waarbij een uitvoeriger audit nodig is. Opvallend is daarbij overigens wel de omgekeerde insteek. Waar risicogericht toezicht in het algemeen is gericht op detectie van onderpresteren (zie hierna), selecteert de door Dixit beschreven methode juist de hoog presterende agenten eerder uit.

Binnen het Nederlandse toezicht heeft het risicogerichte toezicht een min of meer brede, formele status gekregen met de publicatie van de Kaderstellende visie op toezicht (BZK, 2005). In deze Kaderstellende visie werden zes principes van goed toezicht gepresenteerd, waarvan een het principe van selectiviteit was. Selectief toezicht heeft betrekking op twee aspecten, namelijk de mate waarin de overheid zelf voor toezicht zorgdraagt en de vorm en omvang van toezicht op basis van een afweging van risico's, kosten en baten. Voor wat betreft het eerste aspect kan daarbij gekeken worden of vormen van zelfregulering of certificering in voldoende toezicht kunnen voorzien. Het laatste aspect sluit aan bij de omschrijving van risicogericht toezicht zoals gebruikt door Black (2010).

In 2006, in vervolg op de Kaderstellende visie (BZK, 2005), stelde het Beraad van Inspecteurs-Generaal (2006) een leidraad op voor alle inspecties om vorm te geven aan een risicogerichte werkwijze. Daarin

³ Ook risicogestuurd of risicogebaseerd toezicht zijn gangbare termen met dezelfde betekenis.

worden vijf stappen onderscheiden in een afgewogen proces van risicogericht toezicht. Dat zijn achtereenvolgens:

1. Context vaststellen
2. Gevaren identificeren
3. Risico's bepalen en analyseren
4. Toezichtstrategieën
5. Intervenieren

Deze stappen worden omgeven door communicatie en overleg met interne en externe belanghebbenden en door monitoring en evaluatie van de toezichtactiviteiten.

Het begrippenkader van de Rijksinspecties (Velders en Brunia, 2013) geeft geen eenduidige definitie van risicogericht toezicht, maar doet dat wel van de risicoanalyse die ten grondslag ligt aan de inrichting daarvan. De risicoanalyse wordt omschreven als: *"het gestructureerd en gewogen gebruiken van beschikbare kennis om te bepalen hoe vaak gevaren zich kunnen voordoen en hoe groot de gevolgen daarvan kunnen zijn, en het doen van voorstellen hoe de dreiging op te heffen of te beperken met als doel om tot een onderbouwde keuze van prioriteiten in het toezicht te komen."* Dit is een omschrijving die nauw aansluit bij die van Black (2010). Velders en Brunia onderscheiden daarbij een object- en een subjectrisico. Een objectrisico is het risico dat voortvloeit uit de aard van een handeling of zaak zelf. Een subjectrisico vloeit voort uit het nalevingsgedrag van een persoon of organisatie. Het risico bij een object van toezicht (volgens BZK, 2005) wordt daarmee bepaald door een combinatie van het objectrisico en het subjectrisico.

Tijdsvolgordelijk zullen voor een risicoanalyse eerst de objectrisico's worden bepaald. Daarmee worden de algemene prioriteiten voor het toezicht vastgesteld. Een dergelijk proces wordt ook wel aangeduid als programmatisch handhaven (CCV, 2010), hetgeen dan vooral is gericht op het toezicht op de naleving van wettelijke voorschriften. Waar sprake is van kwaliteits- of uitvoeringstoezicht door organisaties met een publieke taak (Velders en Brunia, 2013; BZK, 2005) bestaan die wettelijke voorschriften uit kwaliteitsnormen of randvoorwaarden voor het functioneren van die instellingen. Door middel van een inventarisatie worden die voorschriften of kwaliteitsnormen in beeld gebracht en van prioriteit voorzien waarvan het niet naleven de grootste risico's met zich meebrengt. De in de analyse vastgestelde prioriteit bepaalt in principe de grootte van de handhavinginszets op de te handhaven onderwerpen.

De tweede stap is de analyse van de subjectrisico's. Hierin moet worden bepaald welke onder toezicht staande personen of organisaties de grootste risico's vertonen om tekort te schieten of dat in de toekomst te gaan doen. Bij deze bepaling moet worden nagegaan welke oorzaken er zijn voor het tekortschieten en hoe die oorzaken kunnen worden herkend. Een hulpmiddel hierbij kan een instrument als de Tafel van Elf zijn, dat systematisch nagaat welke motieven en oorzaken er kunnen zijn voor het niet naleven van wettelijke voorschriften (Beraad van Inspecteurs-

Generaal, 2006). De praktische invulling is evenwel sterk afhankelijk van de context waarin elke individuele toezichthouder moet opereren en van welke risicoindicatoren in die context beschikbaar zijn. Een algemeen geldende systematiek is daarvoor niet voorhanden.

1.5 Risicogericht toezicht bij de onderwijsinspectie

Enkele jaren voor het in de Kaderstellende visie op toezicht (BZK, 2005) werd vastgelegd, was, specifiek voor het onderwijs, het principe van selectief toezicht al gecodificeerd bij de vaststelling van de Wet op het onderwijstoezicht (WOT) (Ministerie van OCW, 2002), waarin werd bepaald dat het toezicht op het onderwijs proportioneel zou plaatsvinden. Dit begrip proportioneel betekende dat de inspectie de instellingen niet meer zou belasten dan voor een goede uitoefening van het toezicht noodzakelijk was. Dit kon, volgens de memorie van toelichting bij de WOT in een drietal vormen tot uitdrukking komen, namelijk:

1. in de frequentie van de toezichtonderzoeken;
2. in de intensiteit van de toezichtonderzoeken;
3. in de wijze waarop de inspectie haar informatie voor het toezicht zou vergaren.

In de WOT zelf is dit in eerste instantie uitgewerkt in het voorschrift dat de inspectie op zodanige opereert dat instellingen niet meer worden belast dan voor een zorgvuldige uitoefening van het toezicht noodzakelijk is (art. 4 lid 2, WOT). Bij wijziging van de WOT in 2012 is dit artikel nog aangevuld met een extra lid met het voorschrift dat de intensiteit van het toezicht afhankelijk is van de kwaliteit van het onderwijs, de waarborging van de professionaliteit van instelling en bestuur, de naleving van wettelijke voorschriften en de financiële situatie van de instelling (art. 4 lid 3, WOT). Hiermee is een aantal indicatoren voor risicogericht toezicht, zoals beschreven door Black (2010) ook expliciet in de wet opgenomen.

Zoals Smeets en Verkroost (2011) beschrijven, is de proportionaliteit die de WOT voorschrijft, in de eerste jaren van het toezicht onder de WOT nauwelijks uit de verf gekomen. Dit werd onder meer veroorzaakt door de afspraak met de Tweede Kamer dat alle scholen jaarlijks fysiek door een inspecteur zouden worden bezocht, waardoor de frequentie als selectief mechanisme al niet goed werkte. Dat gold echter ook voor de intensiteit van het toezicht doordat de inspectie er maar moeilijk in slaagde beperkte onderzoeken in te richten. Proportionaliteit bestond vooral uit extra onderzoeksinspanningen. Het derde element van proportionaliteit, informatiebevraging, liep eveneens niet terug omdat scholen niet de veronderstelde zelfevaluaties leverden die daarvoor noodzakelijk waren.

Pas vanaf 2007 kwam de ontwikkeling in een stroomversnelling. Tevoren was de Kaderstellende visie op toezicht verschenen (BZK, 2005) en in vervolg daarop de Visie op Toezicht OCW (OCW, 2006). Daarin waren, naast selectiviteit, ook de andere vijf kenmerken van goed toezicht uit de Kaderstellende visie uitgewerkt, in het bijzonder de kenmerken samenwerkend en slagvaardig. Deze samenwerking heeft vanaf 2008 geleid tot de samenvoeging van de inspectie met de financiële toezichtfunctie van OCW en de toekenning van bevoegdheden tot sanctionering. Dit laatste was tot dan voorbehouden aan de minister. Het doel daarvan was te komen tot een geïntegreerde toezichtuitoefening die gelijktijdig in staat was tekortkomingen bij instellingen zelf te sanctioneren.

De risicogerichte benadering van het toezicht op het onderwijs zelf is vanaf 2007 vrij voortvarend tot stand gekomen (De Wolf en Verkroost, 2011) en is een succesvolle benadering gebleken in het terugdringen van het aantal zeer zwakke scholen en in de tijd die ze erover doen zich te herstellen (Winter, Struiksma en Hollander, 2015; Inspectie van het Onderwijs, 2015). Winter e.a. wijzen daarbij echter ook op enkele neveneffecten van de toezichtbenadering als gevolg van een gebrek aan uitlijning (alignment) (Baker, 2002). Dergelijke neveneffecten worden hierna besproken. De samenvoeging met de financiële toezichtfunctie in 2008 leidde niet direct tot integratie van de toezichtactiviteiten. De evaluatie van de ondergang van Amarantis (Commissie onderzoek financiële problematiek Amarantis, 2012) heeft wel geleid tot ontwikkeling van een geïntegreerde, risicogerichte benadering.

1.6 Effecten en neveneffecten van sturing en toezicht

Sturing en het toezicht op de realisatie door de agent van de met die sturing bedoelde resultaten, kunnen gemakkelijk leiden tot ongewenste effecten. Ordóñez, Schweitzer, Galinsky en Bazerman (2009) beschrijven dat het lange tijd gemeengoed was het stellen van (concrete) doelen te beschouwen als een panacee ter verbetering van de motivatie van werknemers en de prestaties van organisaties. In plaats van vage doelen als "doe zo goed mogelijk je best" zouden concrete, uitdagende doelstellingen de prestaties aanzienlijk verbeteren. Zij betogen echter dat dit stellen van doelen ook een aantal krachtige en voorspelbare neveneffecten kan hebben.

Ordóñez e.a. (2009) geven een aantal voorbeelden van dergelijke afwijkingen. Zo kan er sprake zijn van doelen die te specifiek en te smal zijn. De gerichtheid op die doelen leidt dan tot blindheid voor nevenverschijnselen. Dit is niet zomaar op te lossen door maar een groter aantal doelen te stellen, aangezien men dan geneigd is een aantal doelen

te negeren. De agent selecteert in zo'n geval vooral kwantitatieve boven kwalitatieve doelen en doelen op de korte termijn boven de langere. Ook een onjuiste (te korte) tijdhorizon voor het behalen van de doelen kan negatieve effecten hebben, doordat de agent na het passeren daarvan achterover gaat leunen. Te hoog gestelde doelen kunnen leiden tot het nemen van grotere risico's en onethisch gedrag, zoals bedriegen en het niet (kunnen) halen van die doelen tot demotivatatie. Ordóñez e.a. (2009) vinden verder dat het stellen van doelen het leren en de samenwerking in de weg kunnen staan.

De Wolf en Janssens (2005) spitsen het vraagstuk van ongewenste effecten meer toe op de aansturing van en het toezicht op onderwijsorganisaties. De door hen gesignaleerde verschijnselen zijn te beschouwen als vormen van moral hazard, die volgens Dixit (2002) in de publieke sector het meest voorkomt. Zij onderscheiden allereerst opzettelijk strategisch gedrag, ook wel aangeduid als gaming. Een belangrijke vorm daarvan is window dressing, het creëren van situaties die enkel bedoeld zijn om de toezichthouder gunstig te stemmen. In sommige gevallen kan dat uitlopen in frauduleus gedrag, zoals ook aangeduid door Ordóñez e.a. (2009), bijvoorbeeld door zwakkere leerlingen niet deel te laten nemen aan toetsen of zelfs toetsresultaten te vervalsen. Ook Winter e.a. (2015) signaleren dat scholen op verschillende, oneigenlijke manieren proberen de resultaten te verbeteren, onder meer door extra tijd voor de toetsen te geven of posters met maten en gewichten aan de muur in het examenlokaal te hebben hangen.

Een tweede categorie ongewenste effecten die De Wolf en Janssens onderscheiden, is het onopzettelijk strategisch gedrag. Dat bestaat er in belangrijke mate uit dat de onder toezicht staanden zich onbewust gaan gedragen in een richting die aansluit bij de kijkwijze van de toezichthouder, die tot uitdrukking komt in de gehanteerde prestatiemaatstaven (Baker, 2001). Men kan zich dan richten op procedurele formalisering en "teaching to the test". Ook kan het leiden tot een sterke focus op korte-termijnresultaten. Winter e.a. (2015) merkten dat op bij schoolbesturen die geneigd zijn de kwantitatieve focus op onderwijsopbrengsten te plaatsen boven de meer kwalitatieve doelen die in het onderwijs kunnen worden nagestreefd. Dit sluit aan bij Ordóñez e.a. (2009).

Tenslotte kan toezichtonderzoek leiden tot stress bij de betrokkenen, waardoor de uitkomsten van het onderzoek beïnvloed worden. Winter e.a. (2015) signaleren onder meer dat de methodiek van aankondiging van de toezichtonderzoeken ertoe leidt dat scholen deze uitvoerig voorbereiden en leraren vooraf instrueren. Zelfs, maar dat is dan weer ongewenst strategisch gedrag, mogen mindere leraren soms een dagje thuis blijven tijdens het onderzoek.

De verschijnselen die zich rond sturing en toezicht in het algemeen bij de agent voordoen, kunnen zich in versterkte mate voordoen rond

risicogericht toezicht, en er zijn nog wat extra mogelijkheden. De Wolf en Honingh (2014) noemen van deze laatste onder meer dat risicoindicatoren vaak afgeleide variabelen zijn die worden gebruikt om de werkelijke risico's te detecteren. Dat kan vervolgens leiden tot verschuiving in de perceptie van de betrokkenen, waarbij de risicoindicatoren als de feitelijke risico's worden geïnterpreteerd in plaats van datgene waar zij eigenlijk een indicatie van zouden moeten zijn. In termen van Baker (2002) is er hierdoor sprake van een vergroting van de hoek tussen de doelen van de agent en van de principaal met als gevolg meer ruimte voor ongewenste effecten.

Omdat de toezichthouder risicogericht toezicht toepast om de verificatiekosten (Dixit, 2002) laag te houden (zowel voor zichzelf als voor de onder toezicht staande), is het aantal risicoindicatoren bovendien vaak aanzienlijk kleiner dan het aantal punten waar uiteindelijk op zou worden beoordeeld. Deze risicoindicatoren bepalen vervolgens of een onder toezicht staande wel of niet door een inspectie zal worden bezocht en eventueel aan een verder onderzoek zal worden onderworpen. Verschijnselen van strategisch gedrag zullen zich dan ook in versterkte mate voordoen rond dit kleinere aantal indicatoren (De Wolf en Honingh, 2014). Als het daarmee immers lukt buiten de aandacht van de toezichthouder te blijven, hoeft ook minder inspanning te worden besteed aan de andere beoordelingspunten.

Zo vormen onderwijsopbrengsten een belangrijk onderdeel van de risicodetectie. Als deze onvoldoende zijn (Ehren en Honingh, 2011), leidt dat volgens de risicogerichte systematiek tot een nader onderzoek op de school, waarbij die opbrengsten zelf ook weer onderdeel van de beoordeling zijn. Om te dempen voor incidentele schommelingen komt deze onvoldoende pas tot stand als de opbrengsten drie opeenvolgende jaren onder de norm zijn (Inspectie van het Onderwijs, 2014). Dit kan ertoe leiden dat er een sterk motief is ervoor te zorgen dat in elk geval eenmaal in de drie jaar een resultaat wordt geboekt dat wel aan de norm voldoet en dat strategisch gedrag dus op de loer ligt.

Ehren en Honingh (2011) zijn daarnaast van mening dat het risicomodel van de inspectie suboptimaal werkt. Een belangrijke oorzaak daarvan is dat als belangrijkste risicoindicator de leeropbrengsten worden gebruikt. Zij stellen dat het gebruik van deze indicator scholen aanwijst die reeds disfunctioneren (en waarvan de leerlingen dus al langere tijd slecht onderwijs krijgen) en daarmee dus rijkelijk laat is. Een geschiktere risicoindicator zou scholen moeten aanwijzen (ruim) voordat de resultaten verslechteren zodat nog op tijd kan worden ingegrepen. Veel beter zouden indicatoren uit zelfevaluaties, de kwaliteit van onderwijsprocessen of andere risico's kunnen worden gebruikt die een voorspellende werking hebben, aldus Ehren en Honingh.

Dit is op zichzelf juist en gaat nog voorbij aan het gegeven dat in dit geval de belangrijkste verklarende variabele tegelijk ook de afhankelijke

variabele is. Dat leidt ertoe dat de identificatie van een risico bijna gelijktijdig de vaststelling van een tekortkoming is. Anderzijds kunnen daar enkele kanttekeningen bij worden geplaatst. Allereerst de kanttekening dat voor de invoering van het risicogerichte toezicht door de inspectie feitelijk geen enkele indicator sturend was voor de toezichtinspanningen. Instellingen werden “voor de voet weg” onderzocht en het kon dus veel langer duren voor deze tekortkomingen werden opgemerkt (Smeets en Verkroost, 2011). De risicogerichte benadering leidde in elk geval tot verbetering van de tijdigheid, zij het minder dan theoretisch ideaal zou zijn.

De tweede kanttekening is dat een belangrijke randvoorwaarde bij risicogericht toezicht is dat de toezichthouder efficiënt de toezichtcapaciteit kan toewijzen aan de risico's (Black, 2010) en dat de onder toezicht staanden niet meer dan nodig worden belast uit het oogpunt van proportionaliteit (Smeets en Verkroost, 2011). Alternatieven zoals zelfevaluaties of de analyse van onderwijsprocessen zijn arbeidsintensieve processen. Zij vergen een grote verantwoordingslast door de instellingen of onderzoekslast door de toezichthouder en dragen daarmee niet bij aan de verhoging van de efficiency die met risicogericht toezicht wordt beoogd, maar juist aan verhoging van de verificatiekosten (Dixit, 2002).

1.7 Te onderzoeken vraagstukken

Wil de toezichthouder effectief risicogericht toezicht kunnen houden, dan dient hij te beschikken over een voldoende arsenaal aan indicatoren die iets zeggen over het (toekomstig) functioneren. Hij moet zich daarbij niet te direct focussen op eenvoudige kengetallen die daarmee vaak op korte termijn beïnvloedbaar zijn door opportunistisch gedrag en aldus kunnen bijdragen aan de hiervoor genoemde neveneffecten. Gezien de brede toezichttaak van de inspectie, zowel het toezicht op de resultaten van primaire activiteiten als op de bedrijfseconomische randvoorwaarden, verdient het daarbij aanbeveling dat indicatoren zich over beide terreinen uitstrekken.

In de drie volgende hoofdstukken worden drie van dergelijke benaderingen onderzocht. Bij alle drie geldt als belangrijke randvoorwaarde dat de benodigde informatie om tot de risicoafweging te komen, niet apart hoeft te worden geleverd door schoolbesturen. Dat is primair een politieke beperking van het toezicht: instellingen moeten niet meer worden belast dan voor een goede uitoefening van het toezicht nodig is. Dat zou zich slecht verhouden met informatie die speciaal gemaakt moet worden voor een risicogerichte aanpak. Minstens zo belangrijk is echter dat speciaal vervaardigde informatie al van meet af aan kan lijden onder de hiervoor genoemde problemen van opzettelijk strategisch gedrag. Dat is minder het geval bij informatie die voor andere

doeleinden wordt gegenereerd. Er is dus enkel gebruik gemaakt van informatie die al om andere redenen beschikbaar is gesteld of gekomen.

In hoofdstuk 2 betreft dat informatie die de inspectie zelf heeft verzameld in het kader van de beoordeling van kwaliteitszorg in scholen voor voortgezet onderwijs. Sinds de veranderingen in de aansturing van het onderwijs in de jaren negentig, zoals hiervoor beschreven (Knippenberg en Van der Ham, 1994) is de verantwoordelijkheid voor de kwaliteit van het onderwijs nadrukkelijker gelegd bij de onderwijsinstellingen zelf en worden zij ook geacht zelf systematisch die kwaliteit te borgen en te verbeteren (onder andere artikel 12, lid 1 en 4 WPO en artikel 24, lid 1 en 4 WVO, OCW 1963 en 1981). Sindsdien onderzoekt de inspectie in het kader van haar toezicht de aanwezigheid en de werking van methoden van systematische kwaliteitszorg in de scholen (Smeets en Verkroost, 2011).

Aangezien het uitgangspunt van de kwaliteitszorg is dat het presteren van de scholen erdoor wordt geborgd en verbeterd, zou te verwachten zijn dat die beoordeling van de kwaliteitszorg, die onafhankelijk van de scholen tot stand komt, een voorspellende werking heeft voor de toekomstige kwaliteit en daarmee kan worden gebruikt voor selectie van instellingen in het kader van risicogericht toezicht. In hoofdstuk 2 is de onderzoeksvraag onderzocht in hoeverre dat met de op dat moment gangbare wijze van beoordeling van kwaliteitszorg mogelijk is.

In hoofdstuk 3 staan risico's rond het financieel functioneren centraal. Instellingen zijn ook in financieel opzicht autonoom en verantwoordelijk voor een adequaat financieel beheer. Primair van belang is daarbij de continuïteit van de publieke voorziening van het onderwijs. Risico's voor die continuïteit kunnen een groot negatief effect hebben op de onderwijskwaliteit (Commissie onderzoek financiële problematiek Amarantis, 2012, Commissie onderzoek huisvesting ROC Leiden, 2015).

De Commissie financiële problematiek Amarantis concludeerde dat de bestaande systematiek van de inspectie, detecteren van risico's op basis van kengetallen in de laatste jaarrekening, te betitelen was als toezicht op de achteruitkijkspiegel. Dat is aanleiding geweest tot een aantal maatregelen in de regelgeving. In hoofdstuk 3 wordt onderzocht of een aantal bestaande risicomodellen op basis van financiële ratio's beter in staat is te voorspellen waar zich dergelijke risico's in de nabije toekomst kunnen voordoen. De informatie die hierbij wordt gebruikt is die welke door instellingen regulier in het kader van hun jaarverslaggeving, gecontroleerd door een externe accountant, wordt opgeleverd. Hoewel desondanks niet geheel vrij van mogelijkheden tot strategisch gedrag, zoals ook bij Amarantis duidelijk werd (Commissie onderzoek financiële problematiek Amarantis, 2012), kan deze informatie mogelijk een geschikte basis vormen voor risicodetectie. De onderzoeksvraag is in welke mate dat het geval is.

In hoofdstuk 4 wordt gekeken naar een structuurkenmerk van instellingen dat geheel niet beïnvloedbaar is door strategisch gedrag op het moment van waarneming, namelijk het al dan niet verticaal geïntegreerd zijn van schoolbesturen. Verticale integratie van schoolbesturen is eveneens door de commissie Amarantis geduid als een potentieel risico in verband met de grotere complexiteit waartoe het in de besturing van de instellingen leidt. Anderzijds wordt verticale integratie in de economische theorie ook beschouwd als een mogelijkheid om onzekerheid te reduceren. De onderzoeksvraag is in dit hoofdstuk in hoeverre die verticaliteit als een risico, dan wel als een voordeel te beschouwen is.

Het geheel wordt afgesloten met een aantal conclusies. Daarin zal aandacht worden besteed aan de bruikbaarheid van de onderzochte indicatoren en wordt gekeken naar verdere onderzoeksrichtingen rond de verkenning van nieuwe indicatoren voor het risicogerichte toezicht in het algemeen en speciaal in het onderwijs.

Literatuur

- Beraad van Inspecteurs-Generaal. (2006). In Werkgroep Risicoanalyse (Ed.), *Leidraad objectgericht risicomanagement door rijkstoezichthouders*. Den Haag: Justitie, Expertisecentrum Rechtshandhaving.
- Baker, G. P. (2000). The use of performance measures in incentive contracting. *The American Economic Review*, 90(2), 415-420.
- Baker, G. P., Gibbons, R., & Murphy, K. J. (2002). Relational contracts and the theory of the firm. *Quarterly Journal of Economics*, 117(1), 39-84.
- Black, J. (2010). Risk-based regulation: Choices, practices and lessons being learnt. In OECD (Ed.), *Risk and regulatory policy: Improving the governance of risk* (pp. 185-236) OECD Publishing.
- Black, J., & Baldwin, R. (2010). Really responsive risk-based regulation. *Law and Policy*, 32(2), 181-213.
- Centrum voor Criminaliteitspreventie en Veiligheid (CCV). (2010). *Programmatisch handhaven. gids voor departementen en inspecties*. Utrecht: Het CCV.
- Commissie onderzoek financiële problematiek Amarantis. (2012). *Autonomie verplicht. rapport onderzoek financiële problematiek Amarantis*. Den Haag: Commissie onderzoek financiële problematiek Amarantis.
- Commissie onderzoek huisvesting ROC Leiden. (2015). *Ontspoorde ambitie*. Den Haag: Commissie onderzoek huisvesting ROC Leiden.
- Dixit, A. (2002). Incentives and organizations in the public sector: an interpretative review. *The Journal of Human Resources*, 37(4), 696-727.
- Ehren, M. C. M., & Honingh, M. C. (2011). Risk-based school inspections in the netherlands: A critical reflection on intended effects and causal mechanisms. *Studies in Educational Evaluation*, 37, 239-248.

Elte, R., & Scholtes, E. (2001). *1990-2000 Uit de luwte. Over strategische veranderingen in en rond de onderwijsinspectie*. Utrecht: Inspectie van het Onderwijs.

Gibbons, R. Lecture note 1: Agency theory. Verkregen 01/22, 2016a, van <http://web.mit.edu/rqibbons/www/903%20LN%201%20S10.pdf>

Gibbons, R. Lecture note 2: Relational contracts. Verkregen 22/01, 2016b, van <http://web.mit.edu/rqibbons/www/903%20LN%202%20S10.pdf>

Hendrikse, G. W. J. (2003). Transaction and influence costs. *Economics and management of organizations: Co-ordination, motivation and strategy* (pp. 203-240). London: McGraw-Hill.

Inspectie van het Onderwijs. (2014). *Analyse en waarderings van opbrengsten primair onderwijs*. Utrecht: Inspectie van het onderwijs.

Inspectie van het Onderwijs. (2015). *Zelfevaluatie risicogericht toezicht. Theorie en praktijk van het onderwijstoezicht*. Utrecht: Inspectie van het Onderwijs.

Knippenberg, H., & Ham, W. v. d. (1994). *Een bron van aanhoudende zorg. 75 jaar ministerie van onderwijs [kunsten] en wetenschappen 1918-1993*. Assen: Van Gorcum.

Koninkrijk der Nederlanden. (1815). *Grondwet* [Grondwet voor het Koninkrijk der Nederlanden van 24 augustus 1815]

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK). (2005). *Minder last, meer effect. Zes principes van goed toezicht. Kaderstellende visie op toezicht 2005*. Den Haag: Directie Personeel, Organisatie en Informatie Rijk (POIR).

Ministerie van Onderwijs, Cultuur en Wetenschap (OCW). (1963). *Wet op het voortgezet onderwijs (WVO)* [Wet van 14 februari 1963, tot regeling van het voortgezet onderwijs]

Ministerie van Onderwijs, Cultuur en Wetenschap (OCW). (1981). *Wet op het primair onderwijs (WPO)* [Wet van 2 juli 1981, houdende Wet op het basisonderwijs]

Ministerie van Onderwijs, Cultuur en Wetenschap (OCW). (2002). *Wet op het onderwijstoezicht (WOT)*

Ministerie van Onderwijs, Cultuur en Wetenschap (OCW). (2006). *Toezicht in vertrouwen, vertrouwen in toezicht. Visie op toezicht OCW 2007-2011*. Den Haag: Ministerie van OCW.

Ordóñez, L. D., Schweitzer, M. E., Galinsky, A. D., & Bazerman, M. H. (2009). *Goals gone wild: The systematic side effects of over-prescribing goal setting* (Working Paper 09-083 ed.) Harvard Business School.

Shapiro, S. P. (2005). Agency theory. *Annual Review of Sociology*, 31, 263-284.

Smeets, G., & Verkroost, J. J. H. (2011). *Selectief en slagvaardig. Werken met de WOT (2000-2010)*. Utrecht: Inspectie van het Onderwijs.

Sparrow, M. K. (2000). *The regulatory craft. Controlling risks, solving problems and managing compliance*. Washington D.C.: Brookings Institution Press.

Tirole, J. (1986). Hierarchies and bureaucracies: On the role of collusion in organizations. *Journal of Law, Economics & Organization*, 2(2), 181-214.

Velders, R., & Brunia, M. (2013). *Begrippenkader rijksinspecties*. Den Haag: Inspectieraad.

- Winter, H., Struiksma, N., & Hollander, M. (2015). *De selectiviteit, efficiency en effectiviteit van het risicogerichte toezicht op grond van de WOT*. Groningen: Rijksuniversiteit Groningen.
- Wolf, I. de, & Honingh, M. (2014). Risicogericht toezicht niet vrij van risico's. In F. Mertens, J. Scherpenisse & M. v. d. Steen (Eds.), *Reflecties op de ontwikkeling en professionalisering van het toezicht. 10 jaar leeratelier toezicht en naleving* (pp. 45-58). Den Haag: Nederlandse School voor Openbaar Bestuur.
- Wolf, I. F. de, & Janssens, F. (2005). *Effects and side effects of inspections and accountability in education*. (No. WP53/05). Amsterdam: UvA; Universiteitsbibliotheek.
- Wolf, I. F. de., & Verkroost, J. J. H. (2011). Evaluatie van de theorie en praktijk van het nieuwe onderwijstoezicht. *Tijdschrift Voor Toezicht*, 2(2), 7-24.

2 Voorspelt kwaliteitszorg de prestaties bij instellingen voor voortgezet onderwijs?⁴

Samenvatting

Empirisch onderzoek heeft laten zien dat er samenhangen zijn tussen TQM-indicatoren en prestaties van organisaties. Dit onderzoek heeft vooral betrekking op bedrijven. In dit hoofdstuk wordt bepaald in welke mate het niveau van de kwaliteitszorg in het Nederlandse Voortgezet Onderwijs samenhangt met de prestaties en kan dienen als een risicoindicator voor het (toekomstig) functioneren van de instellingen. Dit is onderzocht bij 130 instellingen in de periode 2006-2013. Daarbij is gebruik gemaakt van primaire data in plaats van de gebruikelijke perceptuele data. De conclusie is dat de huidige wijze van beoordeling van de kwaliteitszorg door de inspectie drie dimensies kent die verschillende verbanden vertonen met het presteren van instellingen. Bij het financieel presteren is dat zeer beperkt het geval, bij onderwijskundig presteren gebeurt dat meer, in het bijzonder waar het gaat om het systematisch analyseren van vorderingen van leerlingen. De bruikbaarheid van de beoordeling van kwaliteitszorg voor risicotaxatie is evenwel beperkt, mogelijk omdat een belangrijke dimensie van kwaliteit, personeelsbeleid, in het beoordelingskader ontbreekt.

2.1 Inleiding

Onderwijsinstellingen⁵ in Nederland zijn zeer autonoom en zijn zelf verantwoordelijk voor hun kwaliteit en continuïteit. Daartoe worden zij geacht systematisch hun kwaliteit te borgen. Het is de algemene opvatting dat een gestructureerde vorm van kwaliteitszorg bijdraagt aan de kwaliteit van het onderwijs. Om die reden let de Inspectie van het Onderwijs

⁴ Dit hoofdstuk is in bewerkte vorm als artikel onder review bij het tijdschrift Mens en Maatschappij.

⁵ Onder een instelling wordt verstaan een schoolbestuur met alle ertoe behorende scholen.

(verder: de inspectie) tijdens haar onderzoeken speciaal op de inrichting en uitvoering van de kwaliteitszorg door de instellingen.

De inspectie verricht risicogericht toezicht naar het presteren van de instellingen. Dat betekent dat zij op basis van risicoanalyse bepaalt welke instellingen een grotere kans lopen op tekortschietende kwaliteit van het onderwijs of problemen op het gebied van de (financiële) continuïteit. Deze instellingen worden bij voorrang onderzocht en zo nodig intervenueert de inspectie vervolgens. Het is in het belang van de instellingen en de inspectie dat deze risicoanalyse zo nauwkeurig mogelijk plaatsvindt met een minimum aan fouten van de eerste en tweede soort. Alle instellingen met risico's moeten worden gedetecteerd, maar zoveel mogelijk moet worden voorkomen dat goed functionerende instellingen ten onrechte gesignaleerd en onderzocht worden.

Het beoordelen van de kwaliteitszorg bij een instelling levert informatie op die uiteraard gebruikt wordt voor de beoordeling van dat moment. Het gegeven echter, dat verondersteld wordt dat de kwaliteitszorg er ook toe bijdraagt dat de huidige en toekomstige kwaliteit en continuïteit van het onderwijs aan de vereisten (blijven) voldoen, zou van die informatie ook een risicoindicator kunnen maken die iets zegt over de te verwachten ontwikkeling van de prestaties. Als voor zo'n bijdrage (empirische) evidentie kan worden gevonden, kan dat de modellen voor de risicoanalyse verder verrijken, waardoor een preciezere detectie mogelijk is. De algemene vraagstelling van dit onderzoek is daarom of de beoordeling van de kwaliteitszorg door de inspectie een voldoende waarborg of indicatie inhoudt van het functioneren van de scholen.

Om deze vraag te onderzoeken, start dit hoofdstuk met een literatuurverkenning van empirisch onderzoek naar de relatie tussen Total Quality Management (TQM) en prestaties. Daarbij wordt de Malcolm Baldrige National Quality Award als referentiemodel gebruikt (National Institute of Standards and Technology (NIST), 1998). Vervolgens wordt ingegaan op de wijze waarop de inspectie kwaliteitszorg bij onderwijsinstellingen beoordeelt en hoe die wijze zich verhoudt tot het Baldrige-model. Dit alles mondt uit in een toespitsing van de vraagstelling en de formulering van een aantal hypothesen voor het onderzoek. Vervolgens wordt beschreven op welke wijze de dataverzameling is samengesteld en welke bewerkingen erop zijn toegepast. Na de beschrijving van de onderzoeksbevindingen wordt een en ander afgesloten met een aantal conclusies en discussie.

2.2 Literatuuronderzoek

2.2.1 Total Quality Management

Powell (1995) en Lee, Zuckweiler en Trimi (2005) beschrijven hoe TQM, waarvan de wortels in Japan liggen, vooral in de jaren tachtig in Amerika opgeld begon te doen. Daaraan lag de vrees ten grondslag dat de Verenigde Staten zouden worden voorbijgelopen door Japan en andere Aziatische landen en tot een tweederangs economische status zouden vervallen. Om die reden nam een aantal vooraanstaande, Amerikaanse bedrijven het voortouw met het invoeren van TQM-methoden. Dat leidde tot een snelle verspreiding, onder meer omdat managers geloofden dat TQM zou bijdragen aan verbetering van de (financiële) resultaten.

In 1987 creëerde de Amerikaanse overheid de Malcolm Baldrige National Quality Award om kwaliteitsverbetering te stimuleren (Dean en Bowen, 1994). Oorspronkelijk alleen voor de industrie en dienstverlenende bedrijven, maar later ook voor gezondheidszorg en onderwijs. Er zijn daarnaast meer methoden ontwikkeld onder de verzamelnaam TQM, waaronder in Europa het EFQM-model. Het Baldrige-model is in de loop van de tijd een soort standaard geworden waar ook in de wetenschappelijke literatuur vaak aan wordt gerefereerd. De belangrijkste domeinen van TQM worden in het Baldrige-model goed weergegeven (Dean en Bowen, 1994).

Het Baldrige-model is ingericht om een TQM-systeem te beoordelen. Het kent een puntensysteem om de positie van organisaties op verschillende aandachtsgebieden te scoren (Lee e.a. 2005). In totaal kunnen 1000 punten worden toegekend, waarbij grofweg een score van 50 procent of hoger per aandachtsgebied als voldoende geldt. De aandachtsgebieden zelf zijn nader onderverdeeld. De aandachtsgebieden in het eerste Baldrige-model van 1988 waren (Lee e.a. 2005):

1. Leiderschap
2. Informatie en analyse
3. Strategische kwaliteitsplanning
4. Personeelsbeleid
5. Kwaliteit(sborging) van producten en diensten
6. Resultaten van de kwaliteitsborging
7. Klanttevredenheid.

Nadien is het model elke vijf jaar herzien. Deze herzieningen zijn voor de periode van 1988 tot en met 2003 beschreven door Lee e.a. (2005). De

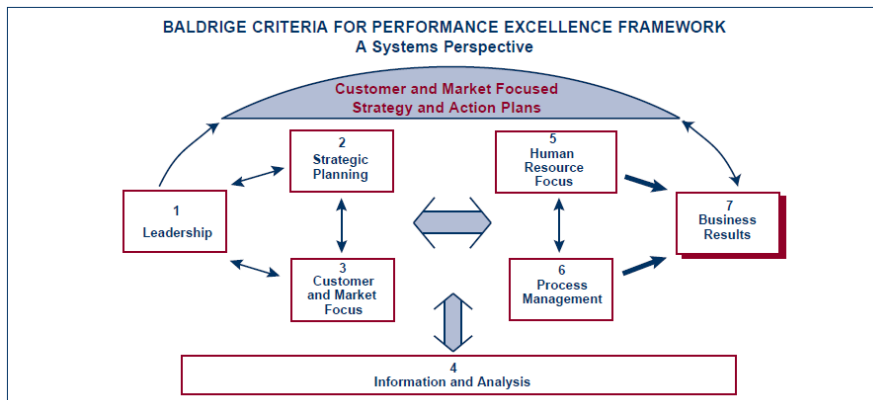
meest ingrijpende verandering was die van 1997, waarbij het aandachtsgebied Klanttevredenheid als belangrijkste doel van de organisatie werd vervangen door resultaten in een veel bredere zin (waaronder financiële). Dit leverde het onderstaande overzicht van aandachtsgebieden op met typeringen:

Tabel 2.1: Typeringen van aandachtsgebieden in het Baldrige-model (NIST, 1998)

	Aandachtsgebied	Typering
1.	Leiderschap	"examines the company's leadership system and senior leaders' personal leadership [...] how senior leaders and the leadership system address values, company directions, performance expectations, a focus on customers and other stakeholders, learning, and innovation"
2.	Strategische planning	"examines how the company sets strategic directions, and how it develops the critical strategies and action plans to support the directions"
3.	Klant- en marktgerichtheid	"examines how the company determines requirements, expectations, and preferences of customers and markets"
4.	Informatie en analyse	"examines the selection, management, and effectiveness of use of information and data to support key company processes and action plans, and the company's performance management system"
5.	Personeelsbeleid	"examines how the company enables employees to develop and utilize their full potential, aligned with the company's objectives"
6.	Procesmanagement	"examines the key aspects of process management, including customer-focused design, product and service delivery, support and supplier and partnering processes involving all work units"
7.	Bedrijfsresultaten	"examines the company's performance and improvement in key business areas — customer satisfaction, financial and

	Aandachtsgebied	Typering
		marketplace performance, human resource results, supplier and partner performance, and operational performance”

Het model gaat ervan uit dat de onderdelen ervan in samenhang van invloed zijn op het uiteindelijke resultaat van de organisatie. Dat is zichtbaar in figuur 3. Het managementproces dat erin wordt weergegeven, loopt in principe van links naar rechts en omvat de aansturing (met de onderdelen leiderschap, strategische planning en klant- en marktgerichtheid) en het resultatenblok (personeelsbeleid, procesmanagement en bedrijfsresultaten). De uiteindelijke focus is gericht op de bedrijfsresultaten, zoals de pijlen weergeven. Dubbele pijlen geven weer dat uitkomsten van onderdelen elkaar wederzijds beïnvloeden.



Figuur 3 Baldrige Criteria for Performance Excellence Framework (NIST, 1998)

2.2.2 Kwaliteitsmanagement en prestaties

Omdat TQM erop gericht is de prestaties van organisaties te verbeteren, is er veel onderzoek gedaan naar de relatie tussen TQM en de prestaties, voor het overgrote deel bij bedrijven. Nair (2006) heeft een metastudie uitgevoerd naar empirische onderzoeken tussen 1995 en 2004 naar deze relatie. Deze onderzoeken hebben bijna allemaal betrekking op “firms” en dan nog meestal “manufacturing plants”. Slechts een onderzoek heeft betrekking op dienstverlenende organisaties, ziekenhuizen (Douglas en Judge, 2001).

Nair (2006) gebruikt in zijn onderzoek de criteria van het Baldrige-model maar wijkt daar op enkele punten van af. Hij onderscheidt zeven dimensies:

- Leiderschap
- Management van mensen
- Procesmanagement
- Productontwerp en -management
- Analyse van kwaliteitsgegevens
- Leveranciersmanagement
- Klantgerichtheid.

Verschillen tussen Nair en het Baldrige-model uit 1997 zitten er onder meer in dat hij strategische planning niet als een afzonderlijke categorie heeft opgenomen. Voorts splitst hij procesmanagement en productontwerp en management, zoals in het model uit 1987. Een apart onderdeel van het procesmanagement, leveranciersmanagement, heeft hij afgezonderd.

Nair (2006) onderscheidt vier performancemaatstaven die uit de verschillende studies naar voren komen, en waaraan hij deze dimensies relateert:

- Financieel (groei in marktaandeel, winstgevendheid, return on assets)
- Operationeel (aspecten van voorraadbeheer)
- Customer service (metingen van klanttevredenheid)
- Productkwaliteit (voldoen aan vereisten, ontwerp kwaliteit, betrouwbaarheid).

Over deze maatstaven heen heeft hij een geaggregeerde maatstaf samengesteld.

Nair komt tot de volgende bevindingen met betrekking tot financiële prestaties en productkwaliteit. Financiële prestaties vertonen positieve correlaties met leiderschap, management van mensen, procesmanagement en klantgerichtheid. De relatie met procesmanagement wordt niet beïnvloed door dempende factoren en is dus vrij direct. In alle andere gevallen zijn er wel dempende factoren. Productkwaliteit vertoont positieve correlaties met leiderschap en klantgerichtheid. Opvallend is dat het niet correleert met productontwerp en -management⁶.

⁶ Nair (2006) spreekt hier ten onrechte over de relatie tussen product design and management en customer service. Hiervan waren, blijkens zijn tabel, onvoldoende gegevens beschikbaar. De gevonden waarde van Ratio1 in zijn artikel komt overeen met de relatie met productkwaliteit.

Een van de onderzoeken waarop Nair (2006) zich heeft gebaseerd, is Kaynak (2003). Zij heeft, met behulp van structural equation modeling, de relaties tussen de verschillende dimensies en prestatie-indicatoren inzichtelijk gemaakt. Zij onderscheidt drie prestatiedimensies: Financial & Market Performance, Quality Performance en Inventory Management Performance. Kaynaks onderzoek laat geen directe relatie zien tussen de kwaliteitszorgdimensies en het financieel presteren. Bij haar loopt deze relatie via Quality Performance. Dit komt voor een belangrijk deel overeen met de bevinding van Nair (2006) die stelt dat er wel een dergelijke relatie is, maar dat die wordt beïnvloed door een aantal dempende factoren. Waar Nair echter stelt dat procesmanagement wel direct zou zijn gerelateerd, laten de resultaten van Kaynak dat niet zien.

Het positieve effect van TQM op financiële prestaties is in de literatuur niet onomstreden. Powell (1995) stelt vast dat grote successen afgewisseld worden met grote mislukkingen. Hij vraagt zich af of TQM kan leiden tot een stabiele, competitieve voorsprong en, als dat zo zou zijn, waar dan de wisselende resultaten en de mislukkingen uit te verklaren zijn. Hijzelf verklaart dit vanuit de resource-based view (o.a. Peteraf, 1993) die stelt dat de heterogeniteit van de hulpbronnen van een bedrijf bepalend is voor de strategie en voor de mogelijkheden een duurzame voorsprong te verwerven op de concurrentie. Daardoor is het moeilijk successen van andere organisaties in de eigen organisatie te imiteren. Dean en Bowen (1994) constateren hetzelfde waar zij stellen dat formulering van de strategie niet enkel moet uitgaan van de verwachtingen van de klanten maar ook van een zorgvuldig onderzoek van de eigen sterkten en zwakten.

Dit sluit aan bij de opvatting van Hendrikse (2004) die stelt dat organisaties te beschouwen zijn als systemen van attributen (resources). Hoewel in principe elke combinatie van attributen in een organisatie denkbaar is, vertoont slechts een beperkt aantal daarvan complementariteit⁷: zij versterken elkaar (of werken elkaar in ieder geval niet tegen). Door het kiezen van een goede combinatie, kan een organisatie in een voordelige positie komen. Implanteren van nieuwe attributen, zoals TQM, zonder daarbij het samenhangend geheel voldoende in de gaten te houden, kan leiden tot het niet realiseren van de gewenste opbrengsten of zelfs tot verliezen.

Deze opvatting komt terug in het onderzoek naar zogenaamde persistent performance differences in seemingly similar enterprises (PPD's) (Gibbons en Henderson, 2013). PPD's zijn het verschijnsel dat organisaties

⁷ Een alternatief begrip hiervoor is synergie.

(bedrijven) die in grotendeels vergelijkbare omstandigheden (waaronder dezelfde markten) opereren, gedurende langere tijd structureel beter, respectievelijk slechter presteren. De verwachting zou zijn dat de gunstige praktijken snel door concurrenten worden overgenomen, waarmee het voordeel verdwijnt, maar dat gebeurt niet. Gibbons en Henderson (2013) leggen mogelijke oorzaken daarvan bij het opereren van het management bij het invoeren van nieuwe managementpraktijken. Zij geven vier mogelijke oorzaken voor het tekortschieten: het probleem niet zien, gebrek aan inspiratie en motivatie en het niet in staat zijn tot implementatie. Vooral dit laatste ligt in lijn met de door Hendrikse (2004) gesignaleerde problematiek rond complementariteiten. De implementatie van TQM is te beschouwen als zo'n managementpraktijk.

Ook vanuit het personeel kan er weerstand ontstaan. Garicano en Rayo (2016) laten aan de hand van principaal-agent-interacties zien dat zo iets het gevolg kan zijn van gevestigde belangen die in het geding komen, maar ook van een behoudende cultuur. Powell (1995) merkt erover op: "But the evidence also suggests that some employees resist or even subvert TQM, finding it ideological or faddish. Furthermore, TQM entails substantial time investment from managers, it is expensive [...], it rarely produces short-term results, it demands intense CEO commitment and it makes unrealistic assumptions about most organizations' capacities to transform their cultures."

Ook het empirisch onderzoek dat relaties heeft gevonden tussen TQM en financiële prestaties is bekritiseerd. York en Miree (2004) hebben gekeken naar de relatie tussen financieel presteren en het winnen van awards zoals Baldrige en vooral naar de gesuggereerde causaliteit. Zij constateerden dat prijswinnaars het inderdaad beter deden, maar dat zij dat tientallen jaren daarvoor ook al deden. York en Miree (2004) draaien de zaak zelfs enigszins om: bedrijven die (financieel) beter presteren kunnen wel eens eerder geneigd zijn hun geld te investeren in TQM.

2.2.3 Kwaliteitsmanagement in het onderwijs

Onderzoek naar de relatie tussen TQM en prestaties in het onderwijs is onder meer gedaan door Badri, Selim, Alshare, Grandon, Younis en Abdulla (2006). Zij baseren zich daarbij op het, enigszins voor de branche aangepaste, Baldrige-model voor het onderwijs (American Society for Quality (ASQ), 2013b). Wat betreft de Organizational performance results onderscheiden Badri e.a. (2006):

- Student learning results: onderwijsopbrengsten

- Student (- and stakeholder – focused) results: klantgerichte oriëntatie, dialoog met omgeving.
- Budgetary, financial and market results: financiële kosten- en risicobeheersing
- Faculty and staff results
- Organizational effectiveness results
- Governance and social responsibility results.

De gevonden relatie tussen process management en performance van Badri e.a. (2006) sluit goed aan bij de bevindingen van Nair (2006) en Kaynak (2003). Opvallend is verder in dit onderzoek dat, in afwijking van het Baldrige-model, Student, stakeholder en market focus niet als een voorwaarde maar als een resultaat, dus als een afhankelijke variabele is opgenomen.

De opkomst van TQM viel gedeeltelijk samen met het proces van autonomievergroting dat zich in het onderwijs vanaf midden jaren tachtig heeft voorgedaan (Knippenberg en Van der Ham, 1994). Dit leidde er onder meer toe dat er meer verantwoordelijkheid voor de onderwijskwaliteit bij de instellingen kwam liggen en dat het (systematisch) zorg dragen voor die kwaliteit een vereiste werd. De vorm waarin instellingen dat doen, is in principe vrij. Voor vo-scholen is zelfs niet expliciet in de wet opgenomen dat zij een kwaliteitszorgsysteem moeten hebben.

2.2.4 Wettelijke eisen aan kwaliteitszorg in het voortgezet onderwijs

De eisen met betrekking tot de zorg voor kwaliteit zijn af te leiden uit twee artikelen in de Wet op het Voortgezet Onderwijs (WVO) (Ministerie van OCW, 1963):

Artikel 23a. Kwaliteit onderwijs

Het bevoegd gezag draagt zorg voor de kwaliteit van het onderwijs op de school. Onder zorg dragen voor de kwaliteit van het onderwijs wordt in elk geval verstaan: het uitvoeren van het in het schoolplan, bedoeld in artikel 24, beschreven beleid op een zodanige wijze dat de wettelijke opdrachten voor het onderwijs en de door het bevoegd gezag in het schoolplan opgenomen eigen opdrachten voor het onderwijs, worden gerealiseerd.

Artikel 24 luidt vervolgens:

Artikel 24. Schoolplan

1. *Het schoolplan bevat een beschrijving van het beleid met betrekking tot de kwaliteit van het onderwijs dat binnen de school wordt gevoerd, en omvat in elk geval het onderwijskundig beleid, het personeelsbeleid en het beleid met betrekking tot de bewaking en verbetering van de kwaliteit van het onderwijs. [...]*

4. *Het beleid met betrekking tot de bewaking en verbetering van de kwaliteit van het onderwijs omvat in elk geval:*

- a. *de wijze waarop het bevoegd gezag bewaakt dat die kwaliteit wordt gerealiseerd,*
- b. *de wijze waarop het bevoegd gezag vaststelt welke maatregelen ter verbetering van de kwaliteit nodig zijn, en*
- c. *maatregelen en instrumenten om te waarborgen dat het personeel zijn bekwaamheid onderhoudt.*

De inspectie heeft dit voorschrift voor de uitvoering van haar toezicht, uitgewerkt in toezichtkaders (Inspectie van het Onderwijs, 2005, 2009, 2012, 2013b). Een toezichtkader omvat in essentie een waarderingskader (waarin de indicatoren, geordend naar kwaliteitsaspecten, beschreven staan waarop de instellingen worden beoordeeld en de normering) en een werkwijze. Zo'n kader, vergelijkbaar met Baldrige, is niet onveranderlijk. Afhankelijk van veranderingen in wetgeving, wetenschappelijke en maatschappelijke inzichten wordt het met regelmaat aangepast.

Ook het kwaliteitsaspect kwaliteitszorg is in de loop van de tijd op een aantal punten aangepast, maar in de kern zijn de belangrijkste elementen voor de beoordeling van de kwaliteitszorg in alle kaders terug te vinden. In het toezichtkader uit 2009 (Inspectie van het Onderwijs, 2009) gaat het om de volgende zeven indicatoren. Al deze indicatoren zijn in alle kaders vanaf 2006 terug te vinden. De beoordelingen vormen gedurende deze hele periode een continue lijn.

d. KWALITEITSZORG

Kwaliteitsaspect 9

De school zorgt systematisch voor behoud of verbetering van de kwaliteit van haar onderwijs.

9.1 De school heeft inzicht in de onderwijsbehoeften van haar leerlingenpopulatie

9.2 De school evalueert jaarlijks de resultaten van de leerlingen

9.3 De school evalueert regelmatig het onderwijsleerproces

9.4 De school werkt planmatig aan verbeteractiviteiten

9.5 De school borgt de kwaliteit van het onderwijsleerproces

9.6 De school waarborgt de kwaliteit van het schoolexamen en van andere toetsinstrumenten

9.7 De school verantwoordt zich aan belanghebbenden over de gerealiseerde onderwijskwaliteit.

In relatie tot de hierboven besproken theorie, moet, voor zover dat al niet duidelijk is geworden uit de beschrijving van de indicatoren, allereerst opgemerkt worden dat de inspectie de kwaliteitszorg niet beoordeelt aan de hand van de normen van TQM. Een belangrijke reden hiervoor is de grondwettelijke vrijheid van inrichting van het onderwijs die Nederland kent (Koninkrijk der Nederlanden, 1815). Het voert te ver daar hier nader op in te gaan, maar het resultaat is dat scholen een grote mate van autonomie hebben bij het inrichten van hun onderwijs en ook van hun kwaliteitszorgsysteem. De wettelijke voorschriften die de verantwoordelijkheid voor de kwaliteitszorg van onderwijsinstellingen bepalen, laten veel ruimte en schrijven TQM niet voor.

TQM biedt evenwel een samenhangend kader en effecten van de elementen uit dat kader zijn al in ruime mate onderzocht (Nair, 2006, Kaynak 2003, Douglas en Judge, 2001, Badri e.a., 2006). Hoezeer ook de instellingen vrij zijn bij de inrichting van hun kwaliteitszorg, ontkomen zij er niet aan volgens bepaalde principes te werken als zij willen voldoen aan artikel 24 lid 4a en 4b van de WVO. De kaders van de inspectie zijn daarop gebaseerd. Door deze te beschouwen in relatie tot een breed en samenhangend kader als TQM, kan aangesloten worden bij bevindingen in bestaand, empirisch onderzoek, zij het dat dat uiteraard slechts voor bepaalde onderdelen kan gelden.

2.2.5 TQM versus de werkwijze van de inspectie

Daar veel onderzoeken naar TQM zijn gebaseerd op het Baldrige-model, ligt het voor de hand de benadering van de inspectie te relateren aan dit model. Omdat veel van het besproken onderzoek rond de eeuwwisseling heeft plaatsgevonden, is hierboven het model uit 1997 (NIST, 1998) als referentiepunt beschreven. In de onderstaande vergelijking van de werkwijze van de inspectie met het Baldrige-model, wordt gebruik gemaakt van de omschrijvingen uit het meest recente model (ASQ, 2013a) en zal de verbinding met het model van 1997 worden aangeduid. Daarnaast zal verwezen worden naar de onderwijsversie van de Malcolm Baldrige National Quality Award (ASQ, 2013b).

Van de zeven dimensies die TQM kent (ASQ, 2013a), zijn er drie waarmee de indicatoren uit het kader van de inspectie te koppelen zijn. Dat zijn:

- Meting, analyse en kennismanagement: indicatoren 9.1, 9.2⁸
- Procesmanagement: indicatoren 9.3, 9.4, 9.5, 9.6
- Klantgerichtheid: indicator 9.7.

Het doel van Meting, analyse en kennismanagement is: *"[...] to guide your organization's process management toward the achievement of key organizational results and strategic objectives, to anticipate and respond to rapid or unexpected organizational or external changes, and to identify best practices to share."* (ASQ, 2013a en 2013b). Kaynak (2003) noemt dit in haar onderzoek Quality data & reporting, Nair (2006) spreekt van Quality data analysis.

Procesmanagement omvat zowel het productieproces als de productvormgeving. De Baldrige-criteria voor onderwijsorganisaties omschrijven het onderdeel werkprocessen (6.1) als: *"[...] the management of your key educational programs and services and work processes, with the aim of creating value for your students and other customers and achieving organizational success and sustainability."* (ASQ, 2013b). Het overeenkomstige item bij de Baldrige-criteria voor bedrijven in het algemeen (die van toepassing zijn op de bedrijven waar het onderzoek van Nair (2006) op gericht is, luidt: *"[...] the management of your key products and work processes, with the aim of creating value for your customers and achieving organizational success and sustainability."* (American Society for Quality, 2013a).

Nair (2006) heeft in zijn onderzoek een splitsing toegepast bij dit item in productontwerp en management van processen. Dat is ook het geval bij Kaynak (2003). Beiden onderzoeken vooral productieondernemingen. Badri e.a. (2006) houden zich strikter aan de criteria met Proces Management. In een organisatie voor (onderwijsdienst)verlening is een scheiding tussen het werkproces en de kwaliteit van de dienst niet goed aan te brengen. Kaynak (2003) merkt dat ook op en heeft daarom de dienstverlenende organisaties buiten haar onderzoek gehouden. Daardoor kan zij in haar analyses een scherpe scheiding tussen management van werkprocessen en productkwaliteit maken, hetgeen bij dienstverlening niet goed mogelijk is.

Het begrip klantgerichtheid kijkt naar: *"[...] how your organization engages its students and other customers for long-term marketplace success, including how your organization listens to the voice of the*

⁸ In 1997 heette dit Informatie en Analyse

customer, builds customer relationships, and uses customer information to improve and to identify opportunities for innovation." (ASQ, 2013b).

Klanten zijn in een organisatie die publieke dienstverlening als onderwijs verzorgt, een lastige kwestie. Verschillende auteurs hebben zich beziggehouden met de vraag wat onder een klant moet worden verstaan. Kwan (1996) maakt een onderscheid tussen externe en interne klanten. Bij externe klanten gaat het om ouders, belastingbetalers en het afnemend bedrijfsleven. Onder interne klanten verstaat Kwan de studenten en de leraren. Deze laatste groep is opvallend, aangezien onderwijspersoneel immers een productiefactor is.

Duidelijk is echter wel dat onderwijs geen heel eenduidig klantbegrip kent. Het begrip "belanghebbenden" dat de inspectie in 9.7 hanteert, weerspiegelt dat. In *Inspectie van het Onderwijs* (2010), een onderzoek in het middelbaar beroepsonderwijs, benoemt zij in elk geval bedrijven, studenten en ouders. Waar een eenduidig klantbegrip ontbreekt, moet de dialoog met belanghebbenden daar een vervanging voor vormen. De ingewikkelde positie van het begrip blijkt ook bij Badri e.a. (2006). Anders dan bij de andere onderzoeken, komt in hun onderzoek "Student, stakeholder and market focus" niet als een exogene (verklarende) variabele voor maar als een (endogene) resultaatvariabele, naast de andere prestaties.

De overige dimensies van het Baldrige-model, leiderschap, strategische planning en personeelsbeleid, ontbreken in het inspectiekader. Dat is opmerkelijk voor het onderdeel personeelsbeleid. Het vloeit immers duidelijk voort uit de wet (WVO artikel 24 lid 4c, Ministerie van OCW, 1963) dat dat onderwerp moet zijn van het kwaliteitsbeleid van instellingen⁹. Het is bovendien een breed gedragen opvatting dat de kwaliteit van het personeel van invloed is op het resultaat van onderwijs. In de aangehaalde onderzoeken is de relatie niet eenduidig. Nair (2006) vindt wel een significant verband tussen personeelsbeleid en financieel presteren maar niet met productkwaliteit. Badri e.a. (2006) vinden een significant verband met presteren in het algemeen.

Waar het betreft de onderdelen leiderschap en strategische planning komt Nair (2006) (die alleen het leiderschap in zijn onderzoek heeft opgenomen) tot de conclusie dat die een duidelijk verband hebben met zowel het financieel presteren als de productkwaliteit, met daartussen wel een aantal dempende factoren die indiceren dat de werking loopt via andere dimensies. Badri e.a. (2006) vinden die werking ook, maar

⁹ In het meest recente toezichtkader is dat inmiddels wel opgenomen, maar dat heeft nog geen resultaten opgeleverd die bruikbaar waren voor dit onderzoek.

signaleren ook een vrij sterke relatie met de organisatieresultaten. Hoe belangrijk het leiderschap ook mag zijn, juist in verband met de vrijheid van inrichting (Koninkrijk der Nederlanden, 1815) is de inspectie zeer terughoudend met het onderzoeken ervan. Het is geen vast onderdeel van het onderzoek naar kwaliteitszorg maar gebeurt alleen bij ernstige problemen.

2.2.6 Kwaliteitszorg en prestaties in het onderwijs in Nederland

Zoals hiervoor duidelijk is geworden, komt het inspectieonderzoek naar kwaliteitszorg op drie terreinen overeen met de categorieën van het Baldrige-model. Op grond van de uitkomsten van de onderzoeken van Nair (2006), Kaynak (2003) en Badri e.a. (2006) zijn de volgende relaties tussen kwaliteitszorg en prestaties te verwachten.

Meting, analyse en kennismanagement (9.1, 9.2)(verderop aangeduid als MAK)

Het analyseren van kwaliteitsgegevens is bij Nair (2006) positief gecorreleerd met de geaggregeerde prestaties van organisaties maar er is onvoldoende ondersteuning voor de correlatie met de afzonderlijke prestatie-indicatoren, in het bijzonder financieel en operationeel. Kaynak (2003) laat in haar model alleen indirecte relaties zien met Quality Performance. Bij Badri e.a. (2006) is er een vrij sterke relatie met de geaggregeerde prestaties.

Procesmanagement (9.3, 9.4, 9.5, 9.6)(verderop aangeduid als PM)

Productontwerp is bij Nair (2006) gecorreleerd met operationele prestaties. Als we kijken naar management van processen (in engere zin volgens Nair), dan is dit bij Nair (2006) positief gecorreleerd met prestaties in het algemeen. Eveneens is er een positieve correlatie met financiële prestaties. Met operationele prestaties was evenwel geen duidelijke correlatie vast te stellen. Daarmee gedragen beide (sub)dimensies zich nagenoeg tegengesteld aan elkaar. Bij Kaynak (2003) die dezelfde splitsing als Nair (2003) hanteert, is er zowel vanuit productontwerp als vanuit management van processen een correlatie met Quality Performance. Deze zijn niet zeer sterk, maar significant. Bij Badri e.a. (2006) is de splitsing van proces management niet gemaakt. Daar is dit gerelateerd aan de geaggregeerde prestaties. Ook hier is het verband niet heel sterk.

Klantgerichtheid (9.7)(verderop aangeduid als KG)

Klantgerichtheid is positief gecorreleerd met de prestaties in het algemeen (Nair, 2006). Dat is het ook met de afzonderlijke prestatie-maten financiën, operationeel, productkwaliteit en customer service (dit laatste lijkt nogal voor de hand te liggen).

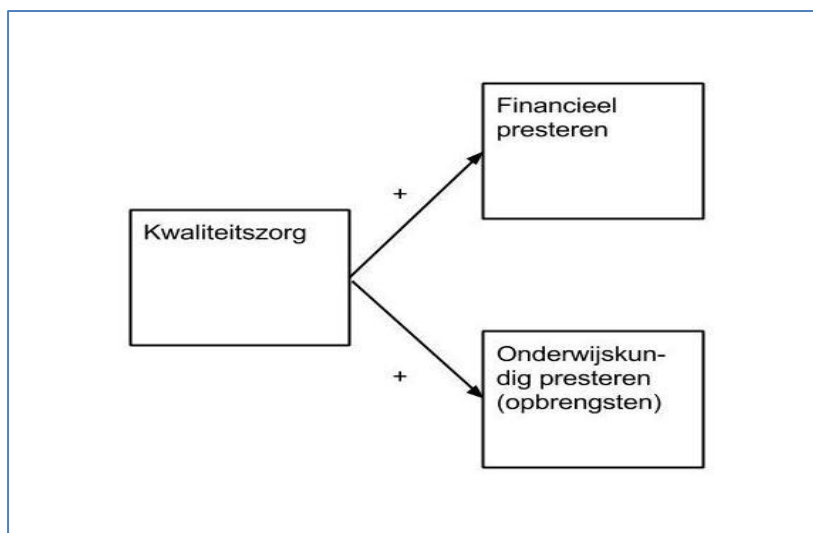
In de aangehaalde onderzoeken (Nair, 2006) zijn de financiële gegevens overwegend verworven als percepties in vragenlijsten. Kaynak (2003) heeft dat eveneens gedaan en heeft afgezien van controle via secundaire data omdat vaak de financiële gegevens op een ander niveau zijn (corporate level) dan waarop het onderzoek zich richt (business unit level). Ook Douglas en Judge (2001) en Badri e.a. (2006) gebruiken alleen gepercipieerde financiële prestaties. Inhoudelijk gezien gaat het daarbij om indicatoren als (groei van) marktaandeel, winstgevendheid en omzet en rentabiliteit op het totaal vermogen (Nair (2006), Kaynak(2003).

Bij onderwijsinstellingen zijn de indicatoren meer gericht op beheersing. Badri e.a. (2006) maken in hun onderzoek in instellingen voor hoger onderwijs gebruik van indicatoren rond kostenbeheersing, efficiënt gebruik van bronnen, scenarioplanning en trendanalyse. Ook hier gaat het om percepties. De beoordeling door de inspectie focust, gezien het risicogerichte karakter van het toezicht, ook vooral op de beheersing van financiële risico's. Dat komt tot uitdrukking in de gekozen indicatoren, gericht op het waarborgen van de continuïteit op korte en langere termijn. De inspectie baseert zich daarbij wel op primaire, financiële gegevens.

Ook de niet-financiële resultaatgegevens zijn bij de onderzoeken van Nair (2006), Kaynak (2003) en Badri (2006) vooral gebaseerd op resultaten uit vragenlijsten. De opbrengstresultaten van de inspectie zijn gebaseerd op primaire gegevens uit het onderwijsproces. Deze worden volgens een ministeriële regeling omgerekend en leiden tot een beoordeling van de opbrengsten per onderwijssoort, per locatie, per jaar (Ministerie van OCW, 2010, Inspectie van het Onderwijs, 2011).

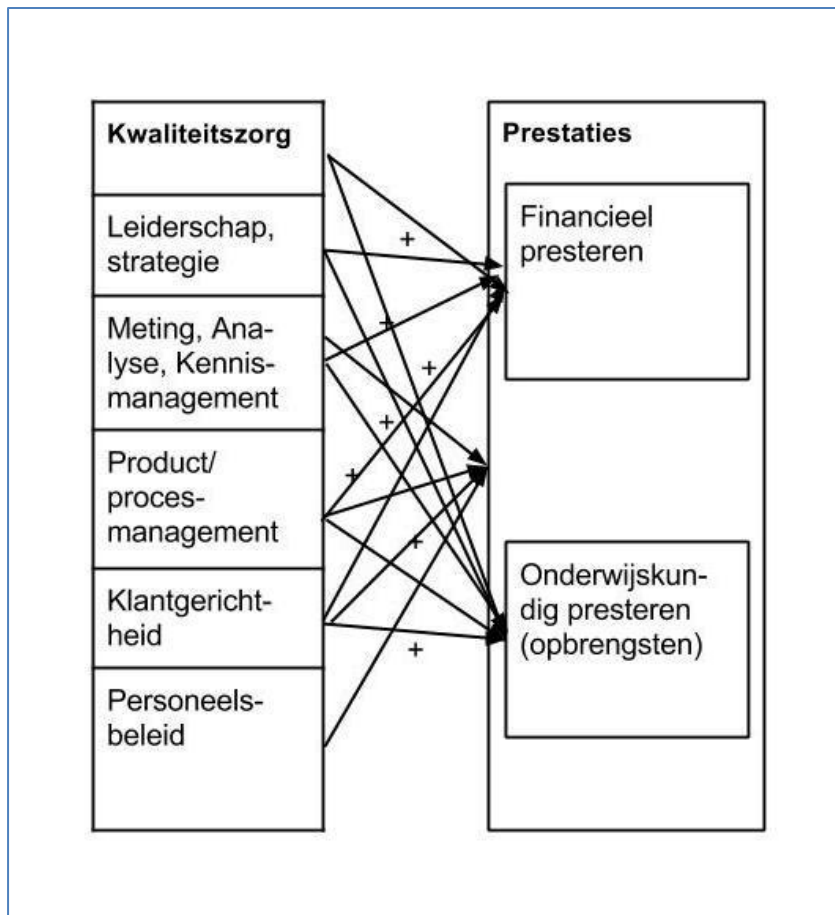
2.3 Vraagstelling en hypothese

Zoals eerder al opgemerkt, is de gedachtegang dat een beoordeling van de kwaliteitszorg door de inspectie een geschikte indicator is van het functioneren op de belangrijkste prestatie-maatstaven, de onderwijskundige en de financiële prestaties. Vertaald naar de vraagstelling van het onderzoek zou het model er in schema als volgt uitzien:



Figuur 4 Relatie kwaliteitszorg en prestaties

De onderzoeken van Nair (2006), Kaynak (2003) en Badri (2006) hebben laten zien dat het onderdeel kwaliteitszorg verder opgesplitst moet worden in verschillende dimensies om de mogelijke relaties beter in beeld te kunnen brengen. Zoals gebleken is, zijn daarbij vijf dimensies van belang: leiderschap, meting, analyse en kennismanagement, procesmanagement, klantgerichtheid en personeelsbeleid. Deze dimensies onderhouden op verschillende manieren relaties met de afzonderlijke of geaggregeerde prestaties, zoals figuur 5 laat zien. Als ze dat met de afzonderlijke prestaties doen, is dat ook met afzonderlijke pijlen weergegeven.



Figuur 5 Relatie kwaliteitszorgdimensies en prestaties

Van deze vijf dimensies van de kwaliteitszorg zijn er drie afzonderlijk terug te vinden in de onderzoeken van de inspectie, meting analyse en kennismanagement (MAK), procesmanagement (PM) en klantgerichtheid (KG). De vraag is of kwaliteitszorg in het algemeen, en in het bijzonder deze onderdelen daarvan, voorspellend zijn voor de onderwijsprestaties en het financieel functioneren van de instellingen. Dat is ook in praktische zin voor een toezichthouder relevant. Als een oordeel over kwaliteitszorg voorspellend zou zijn voor het latere presteren, kan de toezichthouder sneller reageren en met vroege interventies zijn effectiviteit vergroten.

De in de inleiding geformuleerde onderzoeksvraag kan dan ook nader worden gepreciseerd in de volgende vraag:

- In welke mate is de beoordeling van kwaliteitszorg volgens de systematiek van de inspectie verklarend voor het financieel en onderwijskundig presteren van onderwijsinstellingen?

Dat leidt voor het onderzoek tot de volgende, te toetsen hypothesen:

1. MAK is positief gerelateerd aan toekomstig presteren in het algemeen
2. MAK is positief gerelateerd aan toekomstig, onderwijskundig presteren
3. MAK is positief gerelateerd aan toekomstig financieel presteren
4. PM is positief gerelateerd aan toekomstig presteren in het algemeen
5. PM is positief gerelateerd aan toekomstig onderwijskundig presteren
6. PM is positief gerelateerd aan toekomstig financieel presteren
7. KG is positief gerelateerd aan toekomstig presteren in het algemeen
8. KG is positief gerelateerd aan toekomstig onderwijskundig presteren
9. KG is positief gerelateerd aan toekomstig financieel presteren
10. Kwaliteitszorg in het algemeen is positief gerelateerd aan toekomstig onderwijskundig presteren
11. Kwaliteitszorg in het algemeen is positief gerelateerd aan toekomstig financieel presteren
12. Kwaliteitszorg in het algemeen is positief gerelateerd aan toekomstig presteren in het algemeen.

2.4 De onderzoeksgegevens

2.4.1 Instellingen, financiën en onderwijsopbrengsten

Het onderzoek is uitgevoerd op gegevens van bekostigde instellingen voor voortgezet onderwijs in Nederland. Onder een instelling wordt hier verstaan een schoolbestuur met de daartoe behorende scholen en daar weer toe behorende afdelingen. Voor de samenstelling van de onderzoeksgroep is gebruik gemaakt van de volgende gegevens:

- Jaarrekeninggegevens van 2006 tot en met 2012
- Opbrengstoordelen van 2007 tot en met 2013.

Jaarrekeninggegevens worden gepubliceerd op het niveau van het schoolbestuur. Elke instelling heeft dus één jaarrekening. De opbrengstoordelen worden door de inspectie vastgesteld op het laagste, organieke niveau in vo-scholen: de onderwijssoorten per vestiging¹⁰. Een instelling kan, afhankelijk van grootte en complexiteit, van 1 tot n onderwijssoorten hebben.

Alleen die instellingen zijn opgenomen waarvan gegevens over de hele periode beschikbaar waren. Waar fusies plaats hebben gevonden, is per geval beoordeeld of, door consolidatie (zowel financieel als onderwijskundig) van de gegevens van de instellingen voor de fusie, een doorlopende lijn kon worden vastgesteld. In de meeste gevallen kon dat en is het geconsolideerde bestuur over de hele periode opgenomen. In de overige gevallen is het geschrapt. Alleen die instellingen zijn opgenomen die in overwegende mate voortgezet onderwijs uitvoeren. Dit blijkt uit de Jaarrekeninggegevens voor instellingen in het voortgezet onderwijs (Dienst Uitvoering Onderwijs, 2014). Instellingen waar dat niet het geval is, vallen erbuiten en dat geldt ook voor hun opbrengstoordelen.

De opbrengstoordelen komen tot stand volgens een procedure, neergelegd in een ministeriële regeling (Ministerie van OCW, 2010, Inspectie van het Onderwijs, 2011). Indien een onderwijssoort drie jaar lang niet voldoet aan de daarin gestelde normen, wordt het oordeel negatief¹¹. In alle andere gevallen is het positief. Bij onderwijssoorten waar geen opbrengsten zijn vast te stellen, zijn de gegevens buiten het onderzoek gelaten. Waar instellingen uitsluitend praktijkonderwijs verzorgden, waarvoor geen eindexamen geldt en dus niet volgens de systematiek van de inspectie opbrengsten zijn vastgesteld, heeft dat ertoe geleid dat de hele instelling is geschrapt¹².

De bepaling van de uiteindelijke score van een instelling en daarmee de toedeling aan de groep goed dan wel onderpresterende instellingen is gebeurd door indexering van de opbrengstgegevens en de financiële gegevens. De opbrengstbeoordelingen zijn in een aantal stappen geaggregeerd van vestigingen naar scholen naar de instelling en tenslotte

¹⁰ Het gaat daarbij om de volgende onderwijssoorten: vmbo basisberoepsgerichte leerweg (vmbo-bb), vmbo kaderberoepsgerichte leerweg (vmbo-kb), vmbo gemengde en theoretische leerweg (vmbo-gt), havo en vwo.

¹¹ Bij de behandeling van de gegevens zijn de negatieve oordelen als 0 gescoord en de positieve als 1.

¹² In andere gevallen kan het ontbreken van opbrengsten komen doordat een vestiging van een school alleen een onderbouw (juniorcollege) van zo'n onderwijssoort heeft. Gezien de wijze waarop het opbrengstoordeel is opgebouwd, is dat dan niet goed toe te rekenen.

van scores per jaar naar scores over de hele periode door steeds per stap te middelen. Dit leverde voor alle instellingen een score op van 0 tot en met 1.

Voor de financiële scores is gebruik gemaakt van de drie kernindicatoren uit de Beleidsregel financieel toezicht primair en voortgezet onderwijs (Ministerie van OCW, 2011) met de signaleringswaarden:

- Wel/niet een driejarig negatieve rentabiliteit
- Wel/niet een solvabiliteit (incl. voorzieningen) lager dan 0,3
- Wel/niet een liquiditeit (current ratio) lager dan 1.

Hier zijn per indicator waarden tussen 0 en 1 berekend door de scores over de hele periode te middelen. De eindscores per indicator zijn vervolgens met elkaar gemiddeld tot een totaalscore financieel. De geaggregeerde score voor elke instelling is tenslotte berekend door de score voor de opbrengsten te vermenigvuldigen met de score voor de financiën. Dit leverde voor alle instellingen een finale score van 0 tot 1 op. Deze scores waren beschikbaar voor 274 instellingen.

Om te controleren voor andere variabelen die van invloed kunnen zijn op het presteren van de instellingen, is een aantal controlevariabelen toegevoegd. Deze zijn:

- De complexiteit van de instelling. Deze is gebaseerd op de SSG¹³-codes van DUO en onderscheidt instellingen in vijf typen van opklimmende complexiteit. De complexiteit is vormgegeven in vijf dummyvariabelen: SSG1 tot en met SSG5.
- De omvang van de instelling. Deze is gebaseerd op de totale baten van de instelling in het middelste jaar van de analyses, 2009.
- De onderwijskundige, financiële en totale prestaties in het eerste onderzoekstijdvak. Deze prestaties zijn in het tweede tijdvak de afhankelijke variabelen.

2.4.2 De kwaliteitszorgdata

De kwaliteitszorgdata van de inspectie zijn verworven tussen 2006 en 2014. De inspectie beoordeelt kwaliteitszorg op basis van de eerder genoemde indicatoren op een schaal van 1 (slecht) tot 4 (goed). Per onderzoek zijn er maximaal zeven scores, maar niet altijd zijn alle indicatoren gescoord. In principe komt de inspectie tot een eindoordeel op eenzelfde schaal volgens een rekenvoorschrift. In dit onderzoek is het

¹³ De codering van de afdeling Schoolstudiegegevens van DUO.

eendoordeel berekend door de scores van de indicatoren te middelen, waardoor zij een uitkomst opleveren van 1 tot en met 4.

In de loop van de tijd zijn de kaders en de werkwijzen gewijzigd. De belangrijkste breuk in werkwijze heeft plaatsgevonden in 2008. Toen is de inspectie afgestapt van een werkwijze waarbij scholen volgens een vaste ritmiek aan bod kwamen en is zij risicogericht gaan werken. Daarbij werden, op basis van risicoanalyses, scholen met grotere risico's geselecteerd voor de onderzoeken. De onderzochte populatie wordt daardoor minder representatief. Naast deze risicogerichte onderzoeken vonden evenwel ook onderzoeken plaats bij scholen zonder kwaliteitsrisico's, aangezien het parlement bedongen had dat elke school minimaal eenmaal in de vier jaar (licht) moest worden onderzocht en aangezien de inspectie zelf ook een representatief beeld van het onderwijsstelsel moet geven.

Deze verandering kan effect hebben op de beoordeling van de kwaliteitszorg. Te verwachten zou zijn dat de kwaliteitszorg van de instellingen in de loop van de tijd zou verbeteren omdat zij bijleren op dat punt. Gelijktijdig zou te verwachten zijn dat de scores in de risicogerichte onderzoeken vanaf 2008 lager zouden zijn, aangezien het hier gaat om scholen met risico's die mogelijk¹⁴ ook hun kwaliteitszorg minder in orde hebben. Dat zou weer niet gelden voor onderzoeken die uitgevoerd zijn in het kader van de vierjarige beoordeling of in het kader van het representatieve beeld, aangezien daarbij sprake is van instellingen zonder een risicoindicatie. Dit beeld wordt in de scores grotendeels bevestigd. Alleen de laatste reeks onderzoeken laat een terugval in de scores zien.

2.4.3 Bewerking van de gegevens

Om te bepalen of de data afwijkingen vertonen die beperkingen opleggen aan de toetsen of die de uitkomsten beïnvloeden, zijn eerst enkele testen uitgevoerd. Een test op heteroscedasticiteit leverde op dat daar geen sprake van is, noch bij de primaire, noch bij de geaggregeerde gegevens. Een analyse van de residuen versus de voorspelde waarden in een scatterplot laat in alle gevallen een horizontale fit-line zien.

Ook van multicollineariteit is bij de basisdata geen sprake. De variance inflation factor (VIF) is bij de zeven variabelen maximaal 1,695. Dat wordt anders als deze variabelen gecombineerd worden met de samengestelde variabelen (MAK en PM) maar dan is het een artefact van de aggregatie.

¹⁴ Dat is onderwerp van onderzoek.

Bij afzonderlijk beschouwen van de geaggregeerde variabelen komt de VIF niet boven de 1,417. De dummyvariabelen SSG4 en SSG5 vertonen wel multicollineariteit aangezien ze het overgrote deel van de instellingen omvatten en per definitie elkaars tegendeel zijn. Een van beide wordt evenwel steeds weggelaten bij de analyses. Een test op de normaliteit van de gegevens met de Shapiro-Wilks-test laat zien dat geen van de gegevens (noch resultaten, noch kwaliteitszorg) normaal is verdeeld (significantie 0,000). Om die reden is in enkele gevallen gekozen voor non-parametrische toetsen.

Voor de analyses zijn de kwaliteitszorgdata gecombineerd met de resultaatgegevens. Daarbij bleken niet alle instellingen op kwaliteitszorg te zijn onderzocht. Bovendien was het, om de ontwikkeling in de tijd te kunnen onderzoeken, het noodzakelijk de gegevens te splitsen in twee tijdvakken. Dit is gedaan voor het tijdvak ≤ 2009 (2006 tot en met 2009) en ≥ 2010 (2010 tot en met 2013). Deze grens valt in het midden van de te onderzoeken periode. Het was noodzakelijk dat alle instellingen zowel in het eerste als in het tweede tijdvak waren onderzocht op het punt van kwaliteitszorg. Dat leidde tot een reductie van het aantal te onderzoeken instellingen tot 130. Deze groep is op grond van financiën en opbrengsten representatief voor de oorspronkelijke populatie van 274.

De samenstelling van het bestand naar complexiteit ziet er in deze groep van 130 als volgt uit:

- SSG1: alleen vmbo (N=7) (5%)
- SSG2: alleen havo/vwo (N=7) (5%)
- SSG3: vmbo (gt) en havo (evt vwo) (N=13) (10%)
- SSG4: vmbo (breed) en havo (evt vwo) (N=60) (46%)
- SSG5: sectoroverstijgende besturen (N=43) (33%)

De controlevariabele voor de omvang laat zien dat er grote verschillen zijn in de totale baten van de besturen. De gemiddelde totale baten bedragen ruim 35,6 miljoen met een minimum van 1,7 miljoen en een maximum van 461 miljoen. De standaarddeviatie bedraagt 54,2 miljoen. Er is een vrij grote samenhang tussen de complexiteit en de omvang van de instellingen. De correlatie tussen de totale baten en de complexiteit (waarbij de SSG-nummers beschouwd zijn als een ordinale schaal van oplopende complexiteit) bedraagt 0,560 en is significant op 0,01-niveau (1-zijdig, Spearman's rho).

Tabel 2.2 geeft een beschrijving van de centrale variabelen. De tabel laat als meest opvallende verschijnsel zien dat er een vrij forse daling is in de financiële scores ≤ 2009 en ≥ 2010 . Dit komt ook in de andere

onderwijssectoren voor en is een gevolg van de maatregelen van de regering na de bankencrisis.

Tabel 2.2: Beschrijvende statistiek van de belangrijkste variabelen

N=130	Gemiddeld	Standaarddeviatie	Min.	Max.
Financiën vanaf 2010	,7983	,22934	,11	1,00
Financiën tm 2009	,9019	,14544	,33	1,00
Opbrengsten va 2010	,8870	,13552	,38	1,00
Opbrengsten tm 2009	,8896	,13655	,33	1,00
Totaal vanaf 2010	,6982	,23821	,09	1,00
Totaal tm 2009	,7900	,20348	,18	1,00
Kwaliteitszorg va 2010	2,8334	,23768	2,00	4,00
Kwaliteitszorg tm 2009	2,7140	,28843	2,14	3,71

De correlaties tussen de verschillende variabelen, waarbij de kwaliteitszorg tot en met 2009 uitgesplitst is naar de deelcomponenten MAK, PM en KG, zijn weergegeven in tabel 2.3 (waarbij ** significantie aanduidt op 0,01-niveau en * op 0.05-niveau). De tabel laat om te beginnen zien dat er hoge correlaties zijn tussen financiële prestaties in de eerste en de tweede periode en iets lagere maar nog altijd significante correlaties tussen de opbrengsten in die twee opeenvolgende perioden. Kwaliteitszorg en de deelcomponenten daarbinnen vertonen geen relatie met het financieel functioneren, behalve een in omgekeerde richting die niet in de hypothesen is opgenomen. Met de opbrengsten doen ze dat wel, volgens verwachting en de deelcomponenten van kwaliteitszorg blijken ook duidelijk met elkaar en met kwaliteitszorg als geheel te correleren. Dit laatste is enigszins een artefact, aangezien de deelcomponenten zelf uiteraard een onderdeel zijn van kwaliteitszorg in het geheel. Opvallend is dat MAK2009 als enige significant correleert met de kwaliteitszorg vanaf 2010.

Tabel 2.3: Correlaties tussen de variabelen financieel, opbrengsten en kwaliteitszorg

Spearman's rho		Fin2009	Fin2010	Opb2009	Opb2010	Kwa2009	Kwa2010	MAK2009	PM2009
Fin2010	Cor. Coef.	,583**							
	Sig. (2-t.)	,000							
	N	130							
Opb2009	Cor. Coef.	,061	,053						
	Sig. (2-t.)	,488	,552						
	N	130	130						
Opb2010	Cor. Coef.	,171	-,019	,434**					
	Sig. (2-t.)	,051	,829	,000					
	N	130	130	130					
Kwa2009	Cor. Coef.	,018	-,098	,232**	,317**				
	Sig. (2-t.)	,842	,266	,008	,000				
	N	130	130	130	130				
Kwa2010	Cor. Coef.	,211*	,168	-,041	,155	,158			
	Sig. (2-t.)	,016	,056	,641	,079	,072			
	N	130	130	130	130	130			
MAK2009	Cor. Coef.	,036	-,064	,194*	,301**	,635**	,212*		
	Sig. (2-t.)	,682	,469	,027	,001	,000	,016		
	N	129	129	129	129	129	129		
PM2009	Cor. Coef.	-,009	-,126	,264**	,295**	,919**	,119	,417**	
	Sig. (2-t.)	,917	,153	,002	,001	,000	,178	,000	
	N	130	130	130	130	130	130	129	
KG2009	Cor. Coef.	,127	,095	,031	,071	,501**	,099	,242**	,325**
	Sig. (2-t.)	,153	,288	,730	,428	,000	,265	,006	,000
	N	128	128	128	128	128	128	128	128

2.5 Onderzoeksbevindingen

2.5.1 Voorspelling van het financieel presteren

In een eerste regressievergelijking is gekeken naar de voorspelling van het financieel presteren ≥ 2010 door de variabele kwaliteitszorg ≤ 2009 , financiën ≤ 2009 en is gekeken naar de invloed van de controlevariabelen. Tabel 2.4 laat de samenvatting van de uitkomsten hiervan zien. Duidelijk is hieruit dat het financieel presteren in het eerste tijdvak van zeer groot belang is voor het financieel presteren in het tweede tijdvak (de coëfficiënt ervan bedraagt 1,031 in het uitgebreidste model in tabel 2.4). Complexiteit en omvang van de instellingen hebben geen significante invloed, en dat geldt ook voor de beoordeling van kwaliteitszorg in het eerste tijdvak. De coëfficiënt van de kwaliteitszorg is bovendien licht negatief (-0,086 in het eerste model). Hiermee wordt hypothese 11 dus niet ondersteund.

Tabel 2.4: Modelsamenvatting financiën vanaf 2010 met kwaliteitszorg tot en met 2009 en controlevariabelen

Model en voorspellers	R ²	Verandering R ²
1 Kwaliteitszorg ≤ 2009	0,012	,012
2 Model 1 + financiën ≤ 2009	0,442	,431
3 Model 2 + SSG1, SSG2, SSG3 en SSG4	0,452	,010
4 Model 3 + Totale baten	0,452	,000

In de hypothesen is voorts gesteld dat onderdelen MAK, PM en KG van de kwaliteitszorg afzonderlijk van invloed zijn op het financieel presteren. In de onderstaande tabel 2.5 staan de resultaten daarvan weergegeven. Daarbij is het eerder financieel presteren als primaire variabele in het model gehouden, en zijn de overige variabelen afwisselend toegevoegd. KG levert geen significante bijdrage aan de verklaring en dat hypothese 9 dus niet wordt bevestigd. De beide andere leveren afzonderlijk wel een significante bijdrage maar doen dat in de verkeerde richting. De hypothesen 3 en 6 vinden hiermee dus eveneens geen bevestiging.

Tabel 2.5: Vergelijking modellen financiën met dimensies kwaliteitszorg

Variabele	Alleen Fin<=2009	Fin<=2009 met PM	Fin<=2009 met MAK	Fin<=2009 met KG
Intercept	-0,127 (0,097)	0,131 (0,145)	0,144 (0,168)	-0,089 (0,098)
Financiën tot en met 2009	1,026 (0,106)	1,036 (0,104)	1,032 (0,105)	0,982 (0,107)
PM		-0.102 (0,043)		
MAK			-0,095 (0,047)	
KG				0,013 (0,034)
σ^2	0,031	0,030	0,030	0,030
R ²	0,423	0,447	0,444	0,405
N	130	130	129	128

2.5.2 Voorspelling van de opbrengsten

Op analoge wijze als bij het onderzoek naar de financiële prestaties, is in eerste instantie gekeken naar de opbrengsten. Tabel 2.6 laat daar het resultaat van zien. De invloed van kwaliteitszorg is hier duidelijk veel groter dan bij de financiële prestaties het geval is. De invloed van de prestaties in het eerdere tijdvak is weliswaar wat kleiner dan bij de financiële prestaties, maar ook hier zeer dominant. Opvallend is dat de complexiteit in dit geval van significante invloed lijkt te zijn. Dit blijkt het gevolg te zijn van twee uitbijters in de groepen SSG1 en SSG2. Dit zijn kleine groepjes van elk zeven objecten. Weglating van een van de objecten doet het significante effect al verdwijnen. Controle met de omvang van de instellingen (totale baten), die sterk samenhangt met de complexiteit maar over de hele populatie is berekend, laat een verwaarloosbare invloed zien.

Tabel 2.6: Modelsamenvatting opbrengsten vanaf 2010 met kwaliteitszorg tot en met 2009 en controlevariabelen

Model en voorspellers	R ²	Verandering R ²
1 Kwaliteitszorg tm 2009	0,057	,057
2 Model 1 + opbrengsten tm 2009	0,194	,136
3 Model 2 + SSG1 tot en met SSG4	0,285	,091
4 Model 3 + Totale baten	0,287	,002

In de hypothesen is gesteld dat, naast kwaliteitszorg in het algemeen, alle afzonderlijke clusters (MAK, PM, KG) van invloed zijn op de opbrengsten (onderwijskundig presteren). Tabel 2.6 suggereert dat dat voor kwaliteitszorg in het algemeen het geval is. De significantie van kwaliteitszorg als enige variabele bedraagt 0,006 en de coëfficiënt is 0,113. Als echter de zeer dominante factor, de opbrengsten uit het voorafgaand tijdvak, eerst in het model wordt gebracht, blijkt de invloed van kwaliteitszorg een stuk geringer te zijn, zoals tabel 2.7 laat zien. De bijdrage is dan niet meer significant.

Tabel 2.7: Modelsamenvatting opbrengsten vanaf 2010 met opbrengsten tot en met 2009 en kwaliteitszorg tot en met 2009

Model en voorspellers	R ²	Verandering R ²
1 Opbrengsten tm 2009	0,170	,170
2 Model 1 + kwaliteitszorg tm 2009	0,194	,024

Tabel 2.8 laat de relevante verschillen zien tussen dit model inclusief en exclusief de opbrengsten.

Tabel 2.8: Vergelijking modellen kwaliteitszorg met en zonder opbrengsten

Variabele	Uitgebreid model	Beperkt model
Intercept	,351 (,114)	,581 (,110)
Kwazo<=2009	,074 (,038)	,113 (,040)
Opb<=2009	,375 (,081)	
σ^2	0,015	0,017
R ²	,194	,057
n	130	130

Voor de variabele KG geldt dat deze zelfs in zijn eentje, dus zonder correctie voor de opbrengsten uit het eerdere tijdvak, geen significante invloed heeft op de opbrengsten van het latere tijdvak. Hypothese 8 wordt dus niet bevestigd. Bij MAK en PM ligt dat wat anders. Beide hebben een significante invloed op de opbrengsten zonder enige controlevariabele (MAK: sig=0,012, PM: sig=0,006)). Toevoeging van de opbrengsten van het eerste tijdvak aan het model leidt er bij PM toe dat de bijdrage niet meer significant is (sig=0,072), terwijl dat bij MAK nog wel het geval is (sig=0,042). De tabellen 2.9 en 2.10 laten voor respectievelijk MAK en PM de kerngegevens van het uitgebreide en beperkte model zien. Hypothese 5 wordt dus niet ondersteund, hypothese 2 wel.

Tabel 2.9: Vergelijking modellen MAK met en zonder opbrengsten

Variable	Uitgebreid model	Beperkt model
Intercept	,343 (,114)	,615 (,107)
MAK	,070 (,034)	,093 (,036)
Opb<=2009	,382 (,080)	
σ^2	0,015	0,018
R^2	,194	,049
N	129	129

Tabel 2.10: Vergelijking modellen PM met en zonder opbrengsten

Variable	Uitgebreid model	Beperkt model
Intercept	,403 (,097)	,645 (,088)
PM	,057 (,032)	,092 (,033)
Opb<=2009	,374 (,082)	
σ^2	0,015	0,017
R^2	,191	,057
n	130	130

2.5.3 Aanvullende analyses op schoolniveau

Onderzoeken naar kwaliteitszorg vinden in de praktijk op scholen plaats. In het voorafgaande zijn deze geaggregeerd naar besturen. Waar sprake is van een bestuur met slechts een school, is daar geen verschil tussen. Anders is het bij besturen met meer scholen. De beoordeling van de kwaliteitszorg op de scholen en vervolgens de aggregatie daarvan naar het bestuursniveau zouden ertoe kunnen leiden dat daarmee informatie verloren gaat. Om die reden is, alleen voor de relatie tussen kwaliteitszorg en opbrengsten (financiën is niet uit te splitsen naar scholen), een analyse gemaakt op het niveau van de scholen. Daarbij is geen gebruik gemaakt van de controlevariabelen SSG en omvang omdat dat enerzijds niet mogelijk is (het zijn bestuurskenmerken en geen schoolkenmerken) en anderzijds omdat door het opsplitsen van besturen in scholen voor een belangrijk deel al gecorrigeerd wordt voor deze variabelen.

Tabel 2.11 laat zien (in vergelijking met tabel 2.7) dat het weinig verschil uitmaakt of de opbrengsten geanalyseerd worden op bestuurs- of op schoolniveau. Ook hier heeft kwaliteitszorg in zijn eentje een significante invloed op de opbrengsten van het volgende tijdvak, maar daar is geen sprake meer van als de opbrengsten uit het voorafgaand tijdvak in het model worden opgenomen. De significantie daalt dan tot 0,055 bij een coëfficiënt van 0,052, hetgeen betekent dat het nauwelijks nog van invloed is op de uitkomst van de vergelijking. Voor de opbrengsten geldt een significantie van 0,000 en een coëfficiënt van 0,380.

Tabel 2.11: Modelsamenvatting opbrengsten vanaf 2010 met kwaliteitszorg tot en met 2009 op schoolniveau

Model en voorspellers	R ²	Verandering R ²
1 Kwaliteitszorg tm 2009	0,022	,022
2 Model 1 + opbrengsten tm 2009	0,171	,150

Ook voor de variabelen MAK, PM en KG zijn de resultaten volkomen analoog aan die welke zich op bestuursniveau voordoen. KG heeft noch alleen, noch in combinatie met de eerdere opbrengsten een invloed van betekenis. PM heeft dat wel alleen maar niet in combinatie met de eerdere opbrengsten. MAK doet dat in beide gevallen. De uitkomsten zijn op schoolniveau dus niet anders dan ze op bestuursniveau zijn.

2.5.4 Voorspelling van de prestaties in het algemeen

De prestaties in het algemeen zijn samengesteld uit de financiële en de onderwijskundige prestaties door middeling. Tabel 2.12 laat zien dat de kwaliteitszorg uit het voorafgaand tijdvak geen enkele invloed heeft op de aldus samengestelde maat. Ook hier zijn de prestaties uit het verleden dominant. Ook hier lijkt de complexiteit van de instelling op het eerste gezicht een rol van betekenis te spelen, maar het zijn dezelfde twee uitbijters als bij de opbrengsten die het beeld vertekenen. Weglating ervan leidt ertoe dat de significantie van de complexiteit volledig verdwijnt.

Tabel 2.12: Modelsamenvatting totaalresultaten vanaf 2010 met kwaliteitszorg tot en met 2009 en controlevariabelen

Model en voorspellers	R ²	Verandering R ²
1 Kwaliteitszorg tm 2009	0,002	,002
2 Model 1 + opbrengsten tm 2009	0,343	,341
3 Model 2 + SSG1 tot en met SSG4	0,395	,052
4 Model 3 + Totale baten	0,395	,000

Een verdere analyse van de clustervariabelen MAK, PM en KG heeft eveneens geen duidelijke relatie opgeleverd, noch als enige verklarende variabele, noch in combinatie met het resultaat uit de voorafgaande periode. R² is voor al deze variabelen in beide gevallen minder dan 0,01. Om die reden zijn de uitkomsten hier niet meer in tabellen opgenomen en is duidelijk dat de hypothesen 1, 4, 7 en 12 niet ondersteund worden.

2.6 Conclusie en discussie

De beoordeling van kwaliteitszorg door de inspectie vertoont geen (positief) verband met het toekomstig financieel presteren van de instellingen en dat geldt ook voor de te onderscheiden onderdelen. Dat procesmanagement (PM) dat verband niet vertoont, is in overeenstemming met de bevindingen van Nair (2006). Ook klantgerichtheid doet het evenwel niet, hoewel dat volgens Nair verband zou houden met alle resultaatvariabelen, inclusief financiën.

De bestaande methodiek van de inspectie om kwaliteitszorg te beoordelen, is niet ontwikkeld met het oog op financieel functioneren en blijkt dan ook ongeschikt voor de voorspelling ervan. Een zeer dominante factor voor het financieel functioneren in een toekomstige periode, is gelegen in het financieel functioneren in de achterliggende periode. Deze resultaten blijken behoorlijk constant gedurende een langere tijd en geven daarmee een voorbeeld van het bestaan van de zogenaamde PPD's. Deze worden bij financiën kennelijk niet veroorzaakt door de kwaliteitszorg zoals de inspectie die tot dusverre onderzoekt en die vooral gericht is op onderwijskundig functioneren. Meer perspectief zou kunnen liggen in het verder onderzoeken van de bruikbaarheid van financiële voorspellingsmodellen voor onderwijsinstellingen om de precisie ervan te vergroten (onder anderen Altman (1968), Ohlson (1980), Tuckman en Chang (1991) en Gordon, Fischer, Greenlee en Keating (2013). Overigens is daarmee niet meer dan ruim 40 procent van de variantie bij de instellingen verklaard, zodat er nog voldoende ruimte is voor verdere exploratie.

Verder is ook nog de vraag of financieel presteren bij bekostigde (nonprofit) instellingen met een publieke functie wel als een resultaatvariabele moet worden beschouwd, zoals het Baldrige-model doet. Anders dan bij ondernemingen in de markt, is het financiële resultaat hier geen doel. De literatuur laat dat ook wel zien. Waar Nair (2006) bij financiële resultaten spreekt over indicatoren als marktaandeel, winstgevendheid en return on assets, is de benadering bij Badri vooral gericht op kosten- en risicobeheersing. Dat is ook de benadering van de inspectie zelf. Mogelijk moet financieel presteren bij dergelijke organisaties eerder worden gezien als een (proces)voorwaarde dan als een resultaat.

Als het gaat om het onderwijskundig presteren van de scholen is de voorspellingskracht van de opbrengsten uit het recente verleden weliswaar geringer dan in de financiële context, maar ook duidelijk dominant. Daarnaast blijkt meting, analyse en kennismanagement (MAK) positief bij te dragen aan de voorspelling. Kwaliteitszorg in het algemeen en product/procesmanagement (PM) doen dat in zekere mate ook, maar in combinatie met de opbrengsten uit het verleden, is hun bijdrage niet meer significant. Met klantgerichtheid (KG) is er geen duidelijk verband.

De positieve relatie met MAK sluit aan bij de opvatting van de inspectie inzake het zogenaamde opbrengstgericht werken. Dit is "het systematisch, doelgericht werken aan het maximaliseren van prestaties" (Visscher en Ehren, 2011). Aan de basis daarvan ligt het verwerven en systematisch gebruiken van gegevens over de uitgangssituatie en vorderingen van de leerlingen. De uitkomst kan dan ook beïnvloed zijn door het gegeven dat opbrengsten zijn gekozen als prestatie maat voor de instellingen. MAK, zoals beschouwd door de inspectie, is sterk op opbrengsten georiënteerd. Mogelijk zouden beoordelingen van de leskwaliteit beter voorspeld worden door PM en zouden zij dan als intermediaire variabelen te beschouwen zijn.

De prestaties in het algemeen vormen een aggregatie van financieel en onderwijskundig presteren. Deze verschillen qua karakteristiek nogal van elkaar. Hierboven is bovendien al de vraag opgeworpen of financieel presteren wel als resultaatvariabele moet worden beschouwd. Als dat niet het geval is, levert het de vraag op wat de betekenis van een samengestelde prestatie maat, zoals bij Badri e.a. (2006) eigenlijk is. Geen van de veronderstelde verbanden tussen de kwaliteitszorgvariabelen en deze samengestelde prestatie maat is bevestigd.

Opvallend is, zoals opgemerkt, het ontbreken van personeelsbeleid als een onderdeel van kwaliteitszorg in de beoordeling door de inspectie. Het zou te verwachten zijn dat dat, zeker in een dienstverlenende organisatie als het onderwijs, nauw zou samenhangen met de prestaties, in het bijzonder de opbrengsten. Dat dat bij Nair (2006) niet zo duidelijk naar voren komt, kan te maken hebben met het feit dat vrijwel alle onderzoeken die hij bespreekt, uitgevoerd zijn bij productiebedrijven. Badri e.a. (2006), die wel onderwijsorganisaties hebben onderzocht, stellen een significante relatie vast tussen personeelsbeleid en prestaties in het algemeen (zij hebben die niet uitgesplitst).

De gevonden relaties tussen kwaliteitszorg en het presteren van de instellingen zijn dus wat diffuser dan op grond van de literatuur te verwachten zou zijn. Daar kunnen verschillende oorzaken voor zijn. Een verschil met veel onderzoeken is het gebruik van primaire gegevens in plaats van percepties op basis van vragenlijsten. Verder is al opgemerkt dat de inspectie maar een gedeelte van dimensies beoordeelt die ook in het TQM-kader voorkomen. Met name de opname van personeelsbeleid zou kunnen bijdragen aan een verrijking ervan.

Literatuur

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589-609.

American Society for Quality (ASQ). (2013a). *Baldrige criteria for performance excellence* (2013-2014 ed.) American Society for Quality.

American Society for Quality (ASQ). (2013b). *Baldrige education criteria for performance excellence* (2013-2014 ed.) American Society for Quality.

Badri, M. A., Selim, H., Alshare, K., Grandon, E. E., Younis, H., & Abdulla, M. (2006). The Baldrige education criteria for performance excellence framework empirical test and validation. *International Journal of Quality & Reliability Management*, 23(9), 1118-1157.

Dean, J. W. jr, & Bowen, D. E. (1994). Management theory and total quality: Improving research and practice through theory development. *The Academy of Management Review*, 19(3), 392-418.

Dienst Uitvoering Onderwijs. (2014). Open onderwijsdata:

Jaarrekeninggegevens voortgezet onderwijs. Verkregen, 2014, van

http://www.ib-groep.nl/organisatie/open_onderwijsdata/databestanden/vo/Financien/default.asp

Douglas, T. J., & Judge, W. Q. j. (2001). Total quality management implementation and competitive advantage: The role of structural control and exploration. *Academy of Management Journal*, 44(1), 158-169.

Garicano, L., & Rayo, L. (2016). Why organizations fail: Models and cases. *Journal of Economic Literature*, 54(1), 137-192.

Gibbons, R., & Henderson, R. (2013). What do managers do? Exploring persistent performance differences among seemingly similar enterprises . In R. Gibbons, & J. Roberts (Eds.), *The handbook of organizational economics* (1st ed., pp. 680-731). Princeton, New Jersey: Princeton University Press.

Gordon, T. P., Fischer, M., Greenlee, J., & Keating, E. K. (2013). Warning signs: Nonprofit insolvency indicators. *International Research Journal of Applied Finance*, IV(3), 343-378.

Hendrikse, G. W. J. (2004). Organisaties als systemen van attributen. *Tijdschrift Voor Economie En Management*, 49(3), 495-511.

Inspectie van het Onderwijs. (2005). *Toezichtkader VO 2006*. Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs. (2009). *Toezichtkader PO/VO 2009*. Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs. (2010). *Besturing en onderwijskwaliteit in het mbo*. (No. 2010-44). Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs. (2011). *De beoordeling van opbrengsten in het voortgezet onderwijs. hoe werkt het?*. Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs. (2012). *Toezichtkader PO/VO 2012*. Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs. (2013a). *Module kwaliteit bestuurlijk handelen. Beschrijving van indicatoren en aandachtspunten per indicator*.

Inspectie van het Onderwijs. (2013). *Toezichtkader VO 2013*. Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs. (2013c). *De kwaliteit van het bestuurlijk handelen in het funderend onderwijs*. (No. 7). Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs. (2014). *Onderwijsverslag 2012/2013*. Utrecht: Inspectie van het Onderwijs.

Kaynak, H. (2003). The relationship between total quality management practices and their effects on firm performance. *Journal of Operations Management*, (21), 405-435.

Knippenberg, H., & Ham, W. v. d. (1994). *Een bron van aanhoudende zorg. 75 jaar ministerie van onderwijs [kunsten] en wetenschappen 1918-1993*. Assen: Van Gorcum.

Kwan, P. Y. K. (1996). Application of total quality management in education: Retrospect and prospect. *International Journal of Educational Management*, 10(5), 25-35.

- Koninkrijk der Nederlanden. (1815). *Grondwet* [Grondwet voor het Koninkrijk der Nederlanden van 24 augustus 1815]
- Lee, S. M., Zuckweiler, K. M., & Trimi, S. (2005). Modernization of the Malcolm Baldrige national quality award. *International Journal of Production Research*, 44(23), 5089-5106.
- Ministerie van Onderwijs, Cultuur en Wetenschappen. (1963). *Wet op het voortgezet onderwijs (WVO)* [Wet van 14 februari 1963, tot regeling van het voortgezet onderwijs].
- Ministerie van Onderwijs, Cultuur en Wetenschappen. (2010). *Regeling leerresultaten VO*.
- Ministerie van Onderwijs, Cultuur en Wetenschappen. (2011). *Beleidsregel financieel toezicht po en vo 2011*.
- Nair, A. (2006). Meta-analysis of the relationship between quality management practices and firm performance—implications for quality management theory development. *Journal of Operations Management*, 24, 948-975.
- National Institute of Standards and Technology (NIST). (1998). *Malcolm baldrige national quality award. 1998 criteria for performance excellence*. Milwaukee: American Society for Quality.
- Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109-131.
- Peteraf, M. A. (1993). The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal*, 14(3), 179-191.
- Powell, T. C. (1995). Total quality management as competitive advantage: A review and empirical study. *Strategic Management Journal*, 16(1), 15-37.
- Samson, D., & Terziovsky, M. (1999). The relationship between total quality management practices and operational performance. *Journal of Operations Management*, 17, 393-409.
- Tuckman, H. P., & Chang, C. F. (1991). A methodology for measuring the financial vulnerability of charitable nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, 20(4), 445-460.
- Visscher, A., & Ehren, M. (2011). De eenvoud en complexiteit van opbrengstgericht werken. Analyse in opdracht van de kenniskamer van het ministerie van onderwijs, cultuur en wetenschap. Enschede: Vakgroep Onderwijsorganisatie en -management, Universiteit Twente.
- York, K. M., & Miree, C. E. (2004). Causation or covariation: An empirical re-examination of the link between TQM and financial performance. *Journal of Operations Management*, 22, 291-311.

3 Voorspellen van financieel disfunctioneren van scholen

Samenvatting

Recente probleemgevallen hebben de behoefte vergroot tot een betere en eerdere detectie van onderwijsinstellingen die het gevaar lopen in financiële problemen te geraken. Al vanaf de jaren zestig van de vorige eeuw zijn analysemodellen ontwikkeld die op grond van financiële ratio's het risico trachten te voorspellen dat organisaties niet aan hun financiële verplichtingen kunnen voldoen en failliet gaan. Deze modellen, pakketten van financiële indicatoren, zijn vooral toegepast op bedrijven in de marktsector.

In navolging van onderzoeken naar het detecteren van dergelijke problemen bij non-profit-organisaties sinds 1990, wordt in dit hoofdstuk de bruikbaarheid onderzocht van dergelijke modellen voor het voorspellen van financiële problemen bij bekostigde onderwijsinstellingen in Nederland. Dit is gebeurd in verschillende contexten, onder meer met gebruikmaking van verschillende onderscheidingswaarden en een getrapt model van besluitvorming. Belangrijk is daarbij de optimalisatie van fouten van de eerste en tweede soort, tot uitdrukking komend in een maat voor de accuraatheid.

Het onderzoek laat zien dat op basis van de juiste onderscheidingswaarde en via een hiërarchisch besluitvormingsmodel een jaar eerder dan anders het geval zou zijn, een accuraatheid van circa 90 procent te halen in de voorspelling van substantiële financiële risico's.

3.1 Inleiding

3.1.1 Aanleiding

De Inspectie van het Onderwijs houdt integraal toezicht op onderwijsinstellingen. Dit houdt in de praktijk in dat zij kijkt naar het operationeel functioneren van de instellingen, in de praktijk de kwaliteit van het onderwijs, het financieel functioneren, in de praktijk de continuïteit en de rechtmatigheid en op de governance, het bestuurlijk functioneren. Dat doet zij op een risicogestuurde manier. Aan de hand van risicoanalyses tracht zij te detecteren bij welke instellingen een verhoogd

risico bestaat op problemen op deze terreinen. In geval van problemen volgt een nader onderzoek en bij gebleken problemen worden deze onder aangepast toezicht geplaatst.

De risicoanalyses zijn geïntegreerd in de zin dat ze indicatoren bevatten voor alle drie de terreinen en dat deze in samenhang worden beschouwd. Dit hoofdstuk richt zich op onderzoek naar de mogelijke verbetering van de risicoanalyse op een van die terreinen, het financieel functioneren en dan in het bijzonder de continuïteit. De overige dimensies, operationeel presteren, governance en rechtmatigheid blijven daarmee buiten beschouwing.

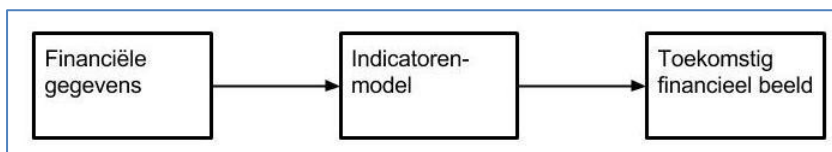
Als instellingen risico's voor de financiële continuïteit vertonen, worden zij dus nader onderzocht en, indien er problemen zijn of dreigen, onder aangepast financieel toezicht (AFT) gezet. De inspectie streeft ernaar de risicodetectie zo precies mogelijk uit te voeren. Zij wil instellingen zonder risico's zo min mogelijk belasten en wil ook zelf zo min mogelijk tijd besteden waar dat niet nodig is. Bovenal wil zij echter niet dat instellingen met risico's over het hoofd worden gezien en dat zij bovendien tijdig worden gedetecteerd, zodat nog voldoende maatregelen mogelijk zijn om het tij te keren.

Zij zoekt dan ook naar methoden om eerder dan tot op heden (waarbij dat in de regel gebeurt na ontvangst van de jaarrekening) indicaties te kunnen krijgen over financiële risico's. Daarbij kunnen voorspellingsmodellen op basis van financiële ratio's een hulpmiddel zijn. Dergelijke modellen zijn al vrij lang in gebruik in de marktsector en de laatste jaren is er ook onderzoek uitgevoerd naar hun toepasbaarheid in de non-profitsector. Het is daarom zinvol te onderzoeken of zij ook toegepast kunnen worden bij het financiële toezicht op het bekostigde onderwijs in Nederland.

De vraag die in het onderzoek dan ook beantwoord moet worden, is op welke wijze de inspectie, met gebruikmaking van ratiomodellen, op een eerder moment optimaal, dus met minimale eigen inspanning en belasting van instellingen en met een maximale detectie van instellingen met risico's, haar toezichttaak kan uitoefenen. In de uitwerking van de vraagstelling zal duidelijk worden dat dit verder gaat dan de simpele vraag om uit een reeks ratiomodellen het beste uit te kiezen, en daarin schuilt ook de wetenschappelijke betekenis van dit onderzoek.

3.1.2 Vraagstelling

De algemene vraag van het onderzoek is dus of en op welke wijze de inspectie in een vroeger stadium financieel risicovolle instellingen kan detecteren. De veronderstelling is daarbij dat toekomstig financieel functioneren, en speciaal disfunctioneren, voor een belangrijk deel te voorspellen is uit financieel functioneren in het recente verleden en dat er indicatoren zijn die een belangrijke bijdrage aan die voorspelling kunnen leveren. Schematisch is dat weer te geven in Figuur 6. Bij de uitwerking van deze vraag zal, zoals in de verdere tekst duidelijk zal worden gemaakt, een aantal modaliteiten nader worden beschouwd.



Figuur 6 Voorspelling op basis van financiële gegevens

Het voorspellen kan volgens verschillende modellen met alle hun eigen karakteristieken. Interessant is daarbij welk model het beste het toekomstig financieel beeld van een onderwijsinstelling kan voorspellen en onder welke condities. Zoals verderop wordt beschreven, is zeer relevant op welke wijze een dusdanig scherp onderscheid wordt gemaakt tussen instellingen met en zonder financiële risico's, dat de inspectie met een minimale belasting van zichzelf en de instellingen tot haar oordelen kan komen. Speciaal is daarbij van belang dat zo min mogelijk risicovolle instellingen over het hoofd worden gezien.

Daar zijn verschillende mogelijkheden voor. Veel van de onderzoeken tot dusverre hebben zich vooral gericht op het analyseren en aanscherpen van een enkel model. Daarbij is van belang hoe een dusdanig onderscheidingswaarde wordt gekozen, dat het aantal type-1- en type-2-fouten minimaal is. In aanvulling op die werkwijze, is het nuttig te bezien in hoeverre combinatie van dergelijke methoden winst op kan leveren. De onderzoeken van Sah en Stiglitz (1986), Hendrikse (1998) en Deng en Hendrikse (2014), die verderop worden besproken, laten zien dat er mogelijk ook winst kan worden geboekt door een afgewogen keuze van opeenvolgende besluitvormingsstappen. Daarbij is het van belang dat de methoden in elke stap uitgaan van een verschillende beoordelingsoptiek.

Dit alles leidt tot de volgende vraagstelling:

Op welke wijze kan een toezichthouder in een eerder stadium in een optimale verhouding van efficiency en effectiviteit financieel risicovolle instellingen voorspellen?

Deze vraag valt uiteen in de volgende deelvragen:

1. Welke modellen op basis van financiële ratio's zijn het meest geschikt voor het voorspellen van het financieel functioneren in een volgend verslagjaar?
2. Leidt een model op basis van het direct voorafgaande jaar tot een beter resultaat dan een model op basis van een eerder jaar?
3. Wat is de optimale onderscheidingswaarde in de logistische regressie?
4. In welke mate leidt een hiërarchisch beslissingsmodel tot een beter resultaat dan de afzonderlijke beslissingsmodellen?
5. Leidt een hiërarchisch model met beslissingseenheden met een verschillende beoordelingsoptiek tot een beter resultaat dan een homogeen model?

3.1.3 Opbouw van het hoofdstuk

Dit hoofdstuk start met een aantal theoretische overwegingen bij verschillende aspecten van het gebruik van ratiomodellen voor de detectie van financiële problemen. Dat betreft allereerst accuraatheid van de modellen en de omgang met fouten van type 1 of type 2. Daarvoor is vervolgens allereerst van belang met welke indicatoren een model is samengesteld en met welke verschillende modellen er in de loop der tijd ervaring is opgedaan. Tenslotte wordt aandacht besteed aan de wijze van besluitvorming. Veel onderzoek naar ratiomodellen is erop gericht het model in een keer een correcte beslissing te laten nemen over de objecten van onderzoek. Onderzoek laat echter zien dat een getrapte besluitvorming vanuit verschillende invalshoeken kan leiden tot een betere besluitvorming in termen van het uitfilteren van risicovolle situaties.

Daarna worden de verschillende modellen besproken die verderop in de analyse zullen worden betrokken. De verschillende indicatoren komen aan de orde en ook komt aan de orde welke aanpassingen noodzakelijk zijn om ze te kunnen gebruiken voor de risicodetectie bij instellingen in het funderend onderwijs. Aansluitend worden de gegevens en de toe te passen methoden van analyse besproken en vinden die analyses plaats.

Het hoofdstuk eindigt met een conclusie en een aantal punten voor verdere discussie.

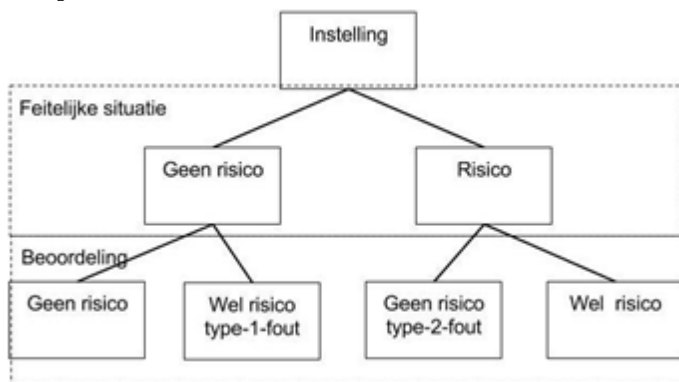
3.2 Theoretische overwegingen

3.2.1 Fouten van type 1 en type 2

Veel onderzoeken naar de betrouwbaarheid van de voorspellingen van verschillende ratiomodellen realiseren percentages correcte classificaties van meer dan 98 procent als het gaat om het risico op insolventie (Gordon e.a. 2013). Een dergelijk percentage is echter vanuit het oogpunt van een toezichthouder weinigzeggend. Dezelfde auteurs (Keating e.a., 2005) wijzen er nadrukkelijk op dat het relevant is een onderscheid te maken naar type-1- en type-2-fouten¹⁵. Zij laten zien dat ook een naïef model, waarin ervan wordt uitgegaan dat geen enkele organisatie insolvent zal raken, al een zeer hoog percentage correcte classificatie haalt, omdat dat maar zo weinig organisaties overkomt. Als er bijvoorbeeld maar twee procent organisaties zijn die insolvent raken, is de succesratio van de voorspelling in dat geval immers 98 procent.

Keating e.a. (2005) stellen dat dit veronderstelt dat de kosten van classificatiefouten gelijk zijn voor beide soorten fouten. Dat is in het algemeen onzeker en Trussel (2002) stelt dat deze kosten lastig te bepalen zijn. Als een bank ten onrechte een lening verstrekt aan een organisatie die die niet kan terugbetalen (type-2-fout), bestaan de kosten

¹⁵ Keating e.a. (2005) spreken van fout-positieven en fout-negatieven. Uitgaande van de hypothese dat het model de riskante objecten correct zal signaleren, zijn fout-positieven de type-2-fouten en fout-negatieven de type-1-fouten. Onderstaand schema laat zien hoe het foutenbegrip in dit hoofdstuk wordt gebruikt



voor de bank uit het verlies van het uitgeleende bedrag. Omgekeerd kan dat ook gelden voor de organisatie als die lening ten onrechte niet wordt verstrekt (type-1-fout), in de vorm van opportunity losses, of tegen aanzienlijk hogere kosten. Wat het kostbaarst is, is niet goed te bepalen. Trussel (2002) onderzoekt deze kosten door verschillende onderscheidingswaarden (ook cut off values genaamd) te hanteren in zijn analyse voor verschillende verhoudingen tussen kosten van fouten van de beide typen.

Voor een toezichthouder zijn er evident verschillende kosten verbonden aan het (niet) reageren op fouten van type-1 of type-2. Als de toezichthouder fouten van type-1 maakt, onderzoekt hij organisaties waar dat niet nodig is. De kosten daarvan bestaan uit de verspilde tijd en bijkomende kosten van zowel de onder toezicht gestelde als de toezichthouder. Type-1-fouten gaan dus ten koste van de efficiency. Bij een fout van type-2 ziet de toezichthouder een risicovolle organisatie over het hoofd. In dat geval bestaan de kosten, naast financieel verlies voor schuldeisers en de overheid, uit reputatieverlies voor de toezichthouder en mogelijk politieke gevolgen. In dat geval is de effectiviteit van de toezichthouder in het geding: hij doet niet goed wat hij moet doen, namelijk tekortschietende instellingen detecteren.

De verdeling van de kansen op wel of niet financieel disfunctioneren, is zeer scheef. Er is een grote groep van organisaties met een kleine kans op disfunctioneren die zich aan een kant van de kansverdeling bevindt, versus het resterende deel dat verspreid ligt over de rest. Om in deze verdeling het beste onderscheid aan te kunnen brengen tussen wel en niet risicovol, waarmee de kans op fouten wordt geminimaliseerd, is het van belang een goede onderscheidingswaarde te kiezen. Greenlee en Trussel (2000) suggereren in hun onderzoek een onderscheidingswaarde van tussen 0,07 en 0,1. Trussel (2002) experimenteert met verschillende onderscheidingswaarden. Gordon e.a. (2013) gebruiken de standaardwaarde van 0,5. Daarmee gaan zij uit van gelijke kosten bij beide typen fouten. Het kiezen van deze onderscheidingswaarde leidt tot een hogere overall score voor het model. Het aantal type-1-fouten wordt erdoor geminimaliseerd, maar dat gaat ten koste van de type-2-fouten. Omdat die laatste maar een zeer kleine fractie van het totaal uitmaken, leidt dat over de totale groep tot een hoge successcore. Als er echter veel belang aan wordt gehecht juist het aantal type-2-fouten te minimaliseren, is een lage onderscheidingswaarde nodig. Dat leidt er dan wel weer toe dat de algemene score van het model daalt. Duda en Schmidt (2010) bespreken verschillende methoden om wat zij noemen de accuraatheid van het model te evalueren. Onder verwijzing naar Jones en Hensher (2008) stellen zij dat het, bij een verdeling waarbij een zeer gering aantal

objecten in de ene of de andere categorie valt, aanbeveling verdient de kansfactor (en daarmee de onderscheidingswaarde) gelijk te stellen aan de verdeling in de populatie.

3.2.2 Keuze van de indicatoren

Indicatoren uit de jaarrekening kunnen worden ontleend aan de balans of de resultatenrekening. Siau (2009) stelt dat er geen goed onderbouwde theorie bestaat voor de weg die organisaties afleggen vanaf het moment dat ze in de problemen beginnen te komen tot het moment dat ze definitief ten onder gaan. Daardoor weet men niet welke indicatoren het zijn die in de loop van dat proces fundamenteel achteruitgaan. Siau stelt vast dat bijna alle studies naar ratiomodellen uitgaan van ratio's die min of meer gemakkelijk beschikbaar zijn of populair, maar die niet op een theorie stoelen. Het model wordt daarmee via trial en error op een empirische wijze tot stand gebracht.

Lev en Sunder (1979) stelden dat dertig jaar eerder ook al vast. Zij geven een aantal theoretische overwegingen bij de keuze van ratio's, maar die zijn voor een meer algemene beoordeling van organisaties of activiteiten en niet speciaal voor het voorspellen van faillissementen. Speciale aandacht besteden zij aan de keuze van het getal in de noemer van de ratio. De belangrijkste reden om gebruik te maken van een bepaalde noemer is corrigeren voor de omvang van de onderzochte variabelen en daarmee de ratio vergelijkbaar te maken met andere waarnemingen. Dit is echter, naar de opvatting van Lev en Sunder aan een aantal voorwaarden gebonden. Een belangrijke daarvan is de onderbouwing van de keuze van de controlevariabele. Lev en Sunder stellen dat die onderbouwing in veel onderzoek ontbreekt en dat de keuze ad hoc is.

De modellen die verderop worden besproken, bevestigen dat beeld voor een belangrijk deel. Voor de keuze van de afzonderlijke ratio's geven alleen Tuckman en Chang (1991) een theoretische motivering. Zij ontleen de ratio's in overwegende mate aan de resultatenrekening. Bij Altman (1968) en Ohlson (1980) hebben empirische overwegingen de belangrijkste rol gespeeld bij de keuze. Zij geven geen theoretische fundering voor hun keuzen. Ohlson stelt zelfs expliciet dat hij geen poging heeft gedaan 'new or exotic' ratio's te ontwikkelen, maar dat het belangrijkste criterium voor de keuze eenvoud was. Bij beide modellen worden de meeste ratio's ontleend aan de balans. Het model van de inspectie is in de loop van de tijd tot stand gekomen en niet in een samenhangende beschrijving theoretisch onderbouwd. Alleen de buffer is

het resultaat van een gericht onderzoek (Don, 2009). Het model kent een mengeling van balans- en resultaatvariabelen.

Lev en Sunder (1979) stellen dat een voorwaarde voor een goede controle op de grootte van de organisatie gelegen is in de proportionaliteit tussen de variabele (de teller) en de controlevariabele. Alleen als daarvan sprake is, biedt de ratio een goede vergelijkingsmogelijkheid tussen de verschillende objecten. Als de correlatie tussen verschillende controlevariabelen voor de omvang van de organisatie hoog is, zijn ze in principe voor dat doel substitueerbaar. Een analyse van de gegevens van de jaarrekening 2012 van alle bekostigde Nederlandse onderwijsinstellingen laat zien dat de correlatie tussen de Totale Activa en de Totale Baten uiterst hoog is (0,953). Waar het gaat om correctie voor de grootte van de instellingen is daarmee de ene niet beter dan de andere.

Intuïtief zou te verwachten zijn dat de keuze voor ratio's die meer op de balans zijn gericht, tot uitdrukking brengt dat ontwikkelingen in de balans van grotere relevantie zijn voor de vraag of de organisatie risico loopt in de problemen te raken. Mutatis mutandis geldt hetzelfde voor een oriëntatie op variabelen uit de resultatenrekening. De te behandelen modellen lijken dat wel enigszins te weerspiegelen. De modellen van Altman (1968) en Ohlson (1980) zijn ontwikkeld vanuit de profitsector. Het model van Tuckman en Chang (1991), als ook het model van de inspectie zijn ontwikkeld vanuit het perspectief van non-profitorganisaties. Het is zeer denkbaar dat een dergelijke achtergrond ook leidt tot een andere invulling van zo'n model met ratio's. Non-profitorganisaties, en zeker zij die voor een belangrijk deel bekostigd worden, zijn voor een veel geringer deel afhankelijk van balanswaarden maar vooral van inkomstenstroom die in de resultatenrekening wordt geregistreerd.

Dat lijkt bevestigd te worden door een onderzoek naar ratio's voor universiteiten door Ritchie en Kolodinsky (2003). Dit heeft een zestal ratio's opgeleverd voor drie dimensies van presteren. Vijf daarvan zijn georiënteerd op de resultatenkant (viermaal kosten en eenmaal baten). De zesde is georiënteerd op de balans (totale activa). Met de verbinding aan dimensies (fundraising efficiency, public support en fiscal performance) brengen zij wel een onderbouwing aan. Niettemin geeft het onderzoek onvoldoende aanwijzingen voor een eenduidige keuze tussen de vier modellen in dit onderzoek. Vast te stellen is dat zij naar hun historie en de keuze van hun controlevariabelen onder te verdelen zijn in twee groepen: meer georiënteerd op de balans versus meer georiënteerd op de resultatenrekening. In het verdere onderzoek zullen ze dan ook naar die twee elementen worden onderscheiden.

Een derde categorie indicatoren zou nog gelegen kunnen zijn in de zogenaamde governance-indicatoren. Dalton, Daily, Ellstrand en Johnson (1998) hebben in een omvangrijk meta-analytisch onderzoek laten zien dat de relatie tussen de samenstelling en organisatie van de ondernemingsleiding en het financieel presteren nauwelijks enige bevestiging vindt. Onderzoek naar oorzaken van falende organisaties (Garicano en Rayo, 2016) levert een aantal cases op waarin de structuur van de besluitvorming leidt tot problemen. De resultaten lenen zich goed voor het voorkomen daarvan in specifieke gevallen, maar leveren geen indicatoren op voor toepassing in een brede risicoanalyse zoals hier wordt beoogd.

Onderzoek naar de relatie tussen de bestuursstructuur en het operationeel presteren bij onderwijsbesturen in het voortgezet onderwijs (Honing, Van Kam en Van Thiel, 2015) vindt vrijwel geen bevestiging. De veronderstelling is dat de afstand tussen enerzijds die structuur en anderzijds het presteren te groot is en dat er teveel dempende factoren tussen zitten. Heemskerk, Heemskerk en Wats (2014) veronderstellen dat eerder gekeken zou moeten worden naar het sociaal handelen en de inspanning van het bestuur en vinden daarin een relatie met het (gepercipieerde) functioneren van het bestuur zelf, maar zij hebben de relatie met het uiteindelijk presteren niet gelegd. Voor de bestuursstructuur van het bekostigd onderwijs is tenslotte sinds 2010 voorgeschreven dat er sprake moet zijn van een scheiding tussen bestuur en (intern) toezicht (Ministerie van OCW, (1981, 1963), Art. 17a en b WPO, Art. 24d en e WVO). Hoewel die scheiding op verschillende manieren vorm kan worden gegeven, ligt het daarom niet voor de hand dat hierin een onderscheidend kenmerk van de besturen kan worden gevonden.

3.2.3 Analysemethoden

Er zijn in de loop van de tijd verschillende methoden ontwikkeld om de ratiomodellen te ontwikkelen en testen. De eerste techniek die onder meer werd toegepast door Altman (1968), betrof de Multiple Discriminant Analyse (MDA). Bij deze techniek worden twee groepen falende en niet falende organisaties paarsgewijs aan elkaar gekoppeld en op een aantal (onafhankelijke) variabelen met elkaar vergeleken (Siau, 2009). Het proces is er vervolgens op gericht de beste onderscheiding tussen beide groepen te vinden. Altman (1968) gebruikte twee groepen van 33 ondernemingen, waarvan de ene groep failliet was gegaan, en baseerde zich bij die groep op het jaarverslag uit het jaar voor het faillissement.

Deze methode leidt tot een formule met als afhankelijke variabele de discriminantscore, door Altman Z genoemd. Deze heeft bij Altman (1968) de volgende vorm:

$$Z = 1.2T_1 + 1.4T_2 + 3.3T_3 + 0.6T_4 + 0.999T_5.$$

De onafhankelijke variabelen T_1 tot en met T_5 worden verderop bij het model van Altman besproken. De Z-score levert een getal op waarbij Altman de kritieke waarde tussen een positieve en negatieve classificatie legde bij 2,67-2,68. Organisaties met een score beneden deze kritieke waarde, worden beschouwd als risicovol. Na toepassing van de formule op een controlegroep rapporteerde Altman 94 procent correcte classificaties.

Mede omdat bij MDA de onafhankelijke variabelen aan vrij strikte normaliteitsvereisten moeten voldoen, pleitte Ohlson (1980) voor de toepassing van logistische regressie en paste deze ook toe voor de ontwikkeling van zijn eigen model. Bij logistische regressie wordt de kans berekend dat een onderneming zal falen gegeven een aantal karakteristieken (Siau, 2009). Ohlson deed dat op basis van een bestand van ruim 2800 bedrijven waarvan er ruim 100 failliet waren gegaan. Ohlson testte zijn model op gegevens een en twee jaar voor het feitelijk faillissement van de ondernemingen. Een jaar voor faillissement bedroeg het percentage correcte classificaties 96 procent.

Tuckman en Chang (1991) hebben voor hun model voor non-profitorganisaties niet van deze technieken gebruik gemaakt. Zij hebben voor hun model vier indicatoren geponeerd en deze toegepast op ruim 4700 non-profit instellingen in verschillende bedrijfscategorieën. Vervolgens hebben zij per indicator het laagste quintiel gekarakteriseerd als "at risk". Organisaties die bij alle indicatoren "at risk" waren, waren "severely at risk". Hun model is op deze wijze door henzelf of anderen niet getoetst op het voorspellend vermogen. Dat is later wel gebeurd door onder meer Greenlee en Trussel (2000) en Trussel (2002), maar dan met behulp van logistische regressie. Hierbij maakten zij gebruik van een controlegroep (hold out sample) waarop zij de gevonden regressievergelijking in de onderzoeksgroep toepasten.

Keating e.a (2005) en Gordon e.a. (2013) betogen, in navolging van Shumway (2001) dat het gebruik van een logistische regressie op basis van slechts een periode tot afwijkende resultaten leidt. Het algemene gebruik bij onderzoekers is voor de analyse van failliete organisaties de jaarrekening te nemen in het laatste jaar voor het faillissement. Voor de overige organisaties is dat jaar evenwel een betrekkelijk willekeurig jaar.

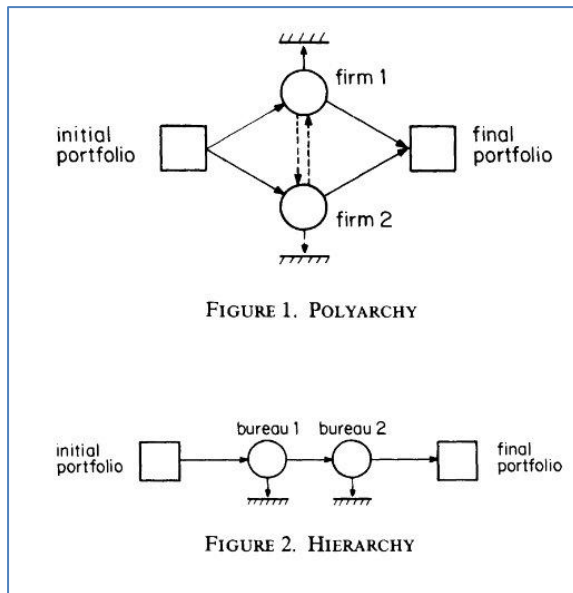
Daarmee zou geen recht worden gedaan aan het gegeven dat tekenen van bankroet zich vaak al veel eerder voordoen en dat zich bovendien, soms per sector, ontwikkelingen in de tijd voordoen die van invloed kunnen zijn. Zij bepleiten daarom een discrete hazard model (discreet tijdmodel (Lammers, Pelzer, Hendrickx en Eisinga, 2007). Daarmee zouden effecten over een langere periode in het regressiemodel worden meegenomen. Gordon e.a. (2013) gebruiken de aldus verkregen modellen op basis van zes methoden van analyse evenwel niet om een voorspelling uit te voeren, maar enkel om de waarde van de modellen zelf te toetsen. Zij hebben dan ook geen controlegroep in hun onderzoek. Hun onderzoek had betrekking op ruim 311.000 non-profitorganisaties in de Verenigde Staten en zij concluderen dat een op Dupont-indicatoren gebaseerd model de gunstigste verhouding heeft tussen voorspelkracht en zuinigheid in het indicatorengebruik.

Verskillende onderzoeken hebben vergelijkingen gemaakt tussen de enkelvoudige logistische methode en het discrete tijdmodel (Rodriguez en Rodriguez, 2006, Fuertes en Kalotychou, 2006, Fantazzini en Figini, 2009, Duda en Schmidt, 2010). Zij komen tot de conclusie dat de verschillen in presteren tussen beide methoden vooral zitten tussen de 'in-sample' en de 'out-of-sample' voorspellingen. Het lijkt er daarmee op dat het gebruik van discrete tijdmodellen geschikter is voor het begrijpen van de mechanismen die (ook over een langere periode) bijdragen aan faillissementen, maar dat voor voorspellingen op de korte termijn de standaard logistische regressie een betere methode is. Gezien de doelstellingen van dit onderzoek, het verkennen van methoden om in een structureel proces sneller te kunnen beschikken over informatie over toekomstig disfunctioneren, wordt bij de analyse dan ook gekozen voor de standaard logistische regressie.

3.2.4 Besluitvorming volgens verschillende processen

Bij alle hierboven besproken methoden wordt gewerkt met een model dat onderscheidt naar wel of niet financieel risicovol. Verbetering van de modellen wordt veelal gezocht in het toevoegen van meer of andere indicatoren of in de analysetechniek. Toevoegen van indicatoren is echter niet onbeperkt mogelijk, aangezien de modellen dan gemakkelijk kunnen gaan lijden onder multicollineariteit. Wat de methoden daarbij buiten beschouwing laten, is dat de keuze tussen wel of niet risicovol, te beschouwen is als een besluitvormingsproces, dat ook op andere manieren kan worden doorlopen. Verschillende soorten besluitvormingsprocessen kunnen leiden tot verschillende uitkomsten en

verschillende maten van effectiviteit. Sah en Stiglitz (1986) onderscheiden polyarchische en hiërarchische besluitvormingsmodellen. Een polyarchisch model is vergelijkbaar met een marktmodel. Parallel aan elkaar nemen organisaties beslissingen over te ondernemen projecten. Een hiërarchisch model is meer vergelijkbaar met een geleide economie of een getrapt model binnen een organisatie. Als de eerste besluitvormer een beslissing heeft genomen, krijgt de tweede enkel nog de projecten die positief zijn bevonden.



Figuur 7 Polyarchie versus hiërarchie (Sah en Stiglitz, 1986)

Sah en Stiglitz (1986) laten zien dat het polyarchische model meer projecten selecteert dan het hiërarchische model. Als het van belang is fouten van de eerste soort te vermijden (ten onrechte verwerpen van goede projecten) werkt het polyarchische model beter. Het hiërarchische model functioneert juist beter waar het gaat om het vermijden van fouten van de tweede soort. Dat is vrij eenvoudig in te zien, zoals ook Sah en Stiglitz betogen, en Figuur 7 laat zien. In de polyarchie wordt een project verworpen als het door beide organisaties wordt verworpen. Als het door beide of door een van beide positief wordt bevonden, kan het worden geaccepteerd. In de hiërarchie wordt een project verworpen als het door een van beide wordt verworpen. Alleen projecten die door beide positief worden bevonden, worden geaccepteerd. Tenzij beide besluitvormingseenheden een in alle opzichten volstrekt identieke kijk op

de werkelijkheid hebben, zullen ze op een aantal punten tot verschillende beoordelingen komen.

Hendrikse (1998) en Deng en Hendrikse (2014) beschrijven in een vergelijking tussen besluitvormingsprocessen van coöperaties en bedrijven die eigendom zijn van investeerders (IOF), de effecten van verschillende besluitvormingsprocessen voor de kans op en de kosten van fouten van de eerste en tweede soort. Bij coöperaties zijn besluitvormingsprocessen vaak getrapt waarbij de eenheden bestaan uit een CEO en een Board of Directors (BoD). De BoD is democratisch samengesteld uit vertegenwoordigers van de leden van de coöperatie. Deng en Hendrikse (2014) gaan er daarbij vanuit dat elk van de opeenvolgende beslissingseenheden vanuit een verschillende beoordelingsoptiek te werk gaat, in het bijzonder bij projecten die vanuit de centrale organisatie worden geïnitieerd (downstream) en projecten die van onderop komen (upstream). Daardoor zullen zij eigenstandig ook verschillende fouten van de eerste en tweede soort maken. Afhankelijk van het type besluiten dat wordt gevraagd en de omgeving, kan een model gekozen worden dat betere resultaten genereert dan de andere.

Hammond (1994) beschrijft hoe de structuur van organisaties van invloed is op de uitkomst van besluitvormingsprocessen. Hij vergelijkt organisaties met functionele structuren en met productgeoriënteerde structuren. Hij laat daarbij zien hoe in een bepaalde structuur gunstige voorstellen verloren gaan in de besluitvormingshiërarchie en hoe minder gunstige voorstellen het juist wel halen. Ook zal de structuur van de organisatie van invloed zijn op de mate waarin de top van de organisatie op de hoogte raakt van vraagstukken (conflicten) die zich op de lagere niveaus afspelen. Om daaraan het hoofd te kunnen bieden, moet de structuur van het accountingsysteem van de organisatie juist de organisatiestructuur doorsnijden. Hendrikse (2004) noemt dat orthogonaliteit. Door in een productgeoriënteerde organisatie een functioneel georiënteerd accountingsysteem te hebben, komt de informatie vanuit twee verschillende invalshoeken bij de top van de organisatie. Dat leidt tot betere informatie aan de top.

De analysemodellen die verderop worden besproken, vertonen verschillen van beoordelingsoptiek. De modellen van Tuckman en Chang (1991) en de inspectie zijn meer georiënteerd op resultaatkant (en in het bijzonder de baten) van de organisatie. Dat ligt voor de hand bij organisaties in de non-profitsfeer die bij hun functioneren minder afhankelijk zijn van duurzame kapitaalgoederen om te produceren, maar meer van een zekere stroom van inkomsten. De modellen van Altman (1968) en Ohlson (1981) zijn veel meer georiënteerd op de activa van de organisatie. Door

dergelijke verschillen in optiek in een getrapt besluitvormingsproces te combineren, kan mogelijk meerwaarde worden verkregen. Het ligt daarbij voor de hand te veronderstellen dat, als er meerwaarde is van een getrapt besluitvormingsproces rond financiële risico's, dit eerder gelegen zal zijn in een combinatie van ongelijksoortige dan van gelijksoortige modellen.

3.2.5 Modellen voor de voorspelling van financiële kwetsbaarheid

Beoordeling en het voorspellen van het financieel opereren van organisaties zijn reeds lang bestaande praktijken in het kader van investeringsbeslissingen en het toekennen van leningen. Op basis van risicoanalyses tracht men de eigen risico's bij dergelijke beslissingen te minimaliseren of bepaalt men welke opslagen noodzakelijk zijn om te compenseren voor die risico's. Het belang van deze risicobeheersing maakt de financiële ratioanalyse al geruime tijd onderwerp van wetenschappelijk onderzoek. Dit onderzoek is gedaan naar onderliggende mechanismen en naar de mogelijkheid om, gebruikmakend van die mechanismen, de toekomstige risico's te voorspellen.

Vanouds heeft de focus van dergelijk onderzoek vooral gelegen op de profitsector (Chabotar, 1989). Markante voorbeelden van dergelijke modellen zijn die van Altman (1968) en Ohlson (1980). Chabotar stelt vast dat de focus bij deze methoden stevast gelegen heeft op return-on-investment en netto winst. Dit zijn ratio's die voor non-profitinstellingen per definitie niet geldig zijn. Enerzijds hebben zij geen aandeelhouders die zij tevreden moeten stellen en anderzijds zijn zij uit hun aard niet op winst gericht. Dat neemt evenwel niet weg, dat verschillende ratio's (ook) nuttig kunnen zijn voor non-profit instellingen en dat er aanvullende ratio's kunnen zijn die speciaal van belang zijn voor de non-profitsector. Voor de bepaling van de risico's van kredietverlening hoeven er niet veel verschillen te zijn tussen bedrijven en non-profitorganisaties (NPO's). Chabotar schetst enkele van die ratio's. Zijn benadering is daarbij voornamelijk gericht op interne verantwoording en sturing.

3.2.6 Modellen voor profitorganisaties

Zoals hierboven is opgemerkt, zijn er twee modellen die brede toepassing hebben in de voorspelling van financiële kwetsbaarheid in de profitsector, de modellen van Altman (1968) en Ohlson (1980). Het model van Altman kent de onderstaande indicatoren. Vooruitlopend op de toepassing van het

model op bekostigde scholen (NPO's) zijn sommige indicatoren van aanvullend commentaar voorzien.

De indicatoren van Altman (1968) zijn::

- WC/TA: working capital/total assets (werkkapitaal/totale active).
- RE/TA: retained earnings/total assets (ingehouden winst/totale activa). Aangezien er in NPO's geen sprake is van dividend op aandelen en dus alle winst wordt ingehouden, moet dit kengetal voor toepassing bij NPO's geherdefinieerd worden. Gordon, Fischer, Greenlee en Keating (2013) gebruiken NA/TA (net assets/total assets, eigen vermogen/totale activa oftewel solvabiliteit) als vervanging. Gezien het feit dat ingehouden winsten toegevoegd worden aan de net assets, is dat een benadering die sterk in het verlengde ligt.
- EBIT/TA: earnings before interest and taxes/total assets (inkomsten voor aftrek van rente en belastingen). EBIT wordt voor scholen berekend door van de nettowinst (voor belastingen) het saldo van financiële baten en lasten af te trekken. Dit komt overeen met het motief van Altman, dat het hier gaat om het verdienvermogen van de productiemiddelen. Winsten uit verdiende renten zijn geen verdiensten uit de primaire activiteiten van scholen.
- MVE/TL: market value of equity/total liabilities (marktwaaarde van de aandelen/totale schulden). Aangezien NPO's geen aandelen uitgeven die verhandeld worden, is geen marktwaaarde te bepalen, anders dan door daar NA voor te nemen, naar analogie van Gordon e.a. (2013). Daarmee krijgt de indicator een zeer grote correlatie met RE/TA als deze geherdefinieerd wordt tot NA/TA.
- S/TA : Sales/total assets (opbrengst verkopen/totale activa). Sales kunnen voor NPO's vertaald worden naar totale baten, dus TR/TA: total revenues/total assets (totale baten/totale activa).

Het model van Ohlson (1980) is, net als dat van Altman, ontwikkeld voor profitorganisaties. Het kent een veel groter aantal variabelen.

- WC/TA: working capital/total assets (identiek aan deze variabele bij Altman)
- CL/CA: current liabilities/current assets (vlottende schulden/vlottende activa) (de inverse van liquiditeit, current ratio)
- NI/TA: net income/total assets (nettowinst/totale activa)
- INTWO: income negative last two years (dichotoom: 1 of 0)
- CHIN: change in net income (in formule: $(NI_t - NI_{t-1}) / (|NI_t| + |NI_{t-1}|)$)

- TL/TA: total liabilities/total assets (totale schulden/totale activa), inverse van solvabiliteit
- OENEG: owner equity negative (negatief eigen vermogen) (dichotoom: 1 of 0), komt overeen met technische insolventie of negatieve solvabiliteit
- SIZE: omvang organisatie gecorrigeerd voor inflatie (in formule: $\ln(\text{total assets}/\text{GNP-price level index})$)
- FFO (Funds from operations)/TL: Inkomsten uit primaire activiteiten/totale schulden. Voor de Nederlandse situatie bij scholen zijn de inkomsten bepaald als de totale baten exclusief rentebaten.

Ohlson (1980) heeft voor zijn model gebruik gemaakt van logistische regressie, terwijl Altman multiple discriminant analyse toepaste. De modellen vertonen inhoudelijk op een aantal punten overlap en zijn overwegend georiënteerd op de activa. In het model van Ohlson zijn enkele ratio's net omgekeerd geformuleerd aan wat gebruikelijk is, maar dat heeft geen invloed op de betekenis ervan. Beide modellen zijn ook in onderzoeken naar non-profitorganisaties toegepast. (onder meer Gordon e.a., 2013; Keating e.a., 2005). Daarvoor waren enkele aanpassingen nodig die hierboven al zijn aangeduid. In deze onderzoeken is het model van Altman (1968) toegepast met behulp van logistische regressie en niet met de originele formule.

3.2.7 Een model voor non-profitorganisaties

Het waren Tuckman en Chang (1991) die een eerste voorzet deden op het terrein van ratioanalyse van non-profitorganisaties met een analysemodel dat beoogde voor een grotere groep van instellingen inzichtelijk te maken welke risico's zij lopen. Zij richtten zich daarbij op het risico dat dergelijke instellingen plotselinge veranderingen in hun (financiële) omgeving, en vooral terugloop in inkomsten, niet op kunnen vangen. Anders dan de hiervoor besproken modellen van Altman en Ohlson, ligt de focus van dit model op de resultatenrekening. Het risico voor de organisaties is er, in de opvatting van de auteurs, in gelegen dat zij door een plotseling terugval in inkomsten gedwongen zijn te snijden in de primaire activiteiten.

Tuckman en Chang (1991) onderscheidden vier indicatoren:

- NA/TR: net assets/total revenues (eigen vermogen/totale baten).
- RCI: revenu concentration index. Deze index geeft weer in welke mate de organisatie afhankelijk is van gespreide inkomstenbronnen (in formule: $RCI = \sum (\text{Revenu}/\text{total revenues})^2$).

Een grotere spreiding zou leiden tot een lagere kwetsbaarheid van de organisatie in geval een van de subsidiënten zou besluiten ermee te stoppen. Voor bekostigde onderwijsinstellingen in Nederland is het de vraag hoe bruikbaar deze indicator is. Deze worden voor het overgrote deel bekostigd door de overheid. De indicator is niet berekend over alle inkomstenbronnen, maar over de vijf rubrieken baten in de resultatenrekening.

- NI/TR: net income/total revenues (nettowinst/totale baten), de rentabiliteit.
- AE/TR¹⁶ : administrative expenses/total revenues. Dit is een wat contra-intuïtieve indicator, aangezien hoge overhead kosten positief meewegen in de analyse. Tuckman en Chang (1991) hebben het risico gedefinieerd in termen van het moeten snijden in de primaire activiteiten. Veel 'vet op de boten' biedt, zeker in de Amerikaanse situatie waar personeel snel kan afvloeien, veel mogelijkheden tot snijden zonder de primaire activiteiten aan te tasten.

Tuckman en Chang (1991) verdeelden, om het risico te bepalen, de te onderzoeken instellingen voor elke indicator onder in quintielen. De instellingen in het ongunstigste quintiel waren "at risk". Instellingen die in alle vier de quintielen onderaan stonden, waren "severely at risk". In hun artikel hebben zij een onderverdeling van instellingen naar dergelijke risico's gemaakt en onderzocht hoe verschillende financiële indicatoren in de aldus onderscheiden groepen uitvielen. Verder hebben zij duidelijk gemaakt dat het noodzakelijk is verschillende soorten non-profitorganisaties van elkaar te onderscheiden. Het maakt bijvoorbeeld nogal uit of er sprake is van een charitatieve organisaties of organisaties in de podiumkunsten. Zij hebben evenwel niet getoetst of het model ook daadwerkelijk goed voorspelde dat organisaties in de problemen zouden komen.

De eersten die een dergelijke toetsing hebben uitgevoerd, waren Greenlee en Trussel (2000). Zij definieerden een financieel kwetsbare organisatie als een die in drie opeenvolgende jaren haar programmakosten als

¹⁶ Administratiekosten/totale baten. Gezien de percentages van 25-30 procent in Tuckman en Chang (1991) dient dit begrip vermoedelijk vertaald te worden met het bredere begrip overheadkosten (administration betekent in het Engels zowel administratie als directie of bestuur). Het begrip roept ook elders in de literatuur verwarring op. Greenlee en Trussel (2000) komen in hun onderzoek uit op 14 procent, terwijl Hager (2001) op vier procent uitkomt. Trussel (2002) gebruikt de indicator niet omdat deze niet in de data voorkomt die hij voor dat onderzoek ter beschikking heeft. De jaarrekeningen van Nederlandse onderwijsinstellingen kennen wel een post administratieve kosten, maar die is van nogal beperkte omvang, namelijk rond vier procent en dekt zeker niet het totaal aan overheadkosten.

percentage van de totale baten had gereduceerd. Dit sluit aan bij de benadering van Tuckman en Chang (1991) die het risico omschreven als de kans te moeten snijden in de primaire activiteiten. In hun onderzoek kwam dat erop neer dat zij dit bepaalden voor de jaren 1993 tot en met 1995 en vervolgens onderzochten of zij de financiële positie in 1995 met behulp van de vier indicatoren konden voorspellen op basis van de jaarrekening over 1992. Zij gebruikten daarbij niet de quintielenindeling van Tuckman en Chang maar een logistisch regressiemodel en onderzochten verschillende onderscheidingswaarden. Over het geheel genomen realiseerden zij een hoogste percentage correcte voorspellingen van 77.

Hager (2001) deed een vervolgonderzoek onder organisaties in de kunstensector. Hij definieerde een kwetsbare organisatie als een die in vier opeenvolgende jaren (1994 tot en met 1997) geen enkele keer gerapporteerd had aan de Internal Revenue Service (de Amerikaanse belastingdienst) en dat in een van de jaren 1990 tot en met 1992 wel had gedaan. Op grond daarvan concludeerde hij dat de organisatie was verdwenen, aangezien vaak geen sprake is van een formeel faillissement. Op basis van een gemiddelde score over 1990 tot en met 1992 heeft Hager getracht het verdwijnen van organisaties te voorspellen met de Tuckman en Chang indicatoren. Het resultaat daarvan was nogal wisselend. Bij veel soorten organisaties leverden zij geen significante resultaten en soms gold dat alleen voor enkele van de indicatoren.

Bij latere onderzoeken (o.a. Gordon e.a., 2013) is de financiële kwetsbaarheid ook wel geoperationaliseerd als de kans op technische insolventie in een van de vervolgjaren. Gordon e.a. (2013) maken voorts een onderscheid naar verschillende soorten instellingen, waaronder onderwijs omdat de verschillende karakteristieken van die organisaties ook verschillend kunnen uitwerken op hun kwetsbaarheid en het belang van bepaalde kengetallen (Hager, 2001).

3.2.8 Het toezichtmodel van de onderwijsinspectie

Het financiële toezicht op onderwijsinstellingen in Nederland heeft zich de afgelopen 25 jaar ontwikkeld van voornamelijk een oriëntatie op rechtmatigheid naar een oriëntatie die steeds sterker gericht is op continuïteit. Dat is een gevolg van de toegenomen autonomie die instellingen hebben gekregen en die in belangrijke mate gestalte heeft gekregen in de zogenaamde lumpsum-bekostiging (Knippenberg en V.d. Ham, 1994). Voordien werden instellingen bekostigd op declaratiebasis.

Volgens een zeer gedetailleerd stelsel van voorschriften ontvingen zij geld voor specifieke uitgaven. Zo werden personeelsleden op de persoon aan de instelling betaald door het ministerie, waarna de instelling dat enkel weer doorsluisde aan het personeelslid. Onder die omstandigheid was het aantal vrijheidsgraden voor instellingen uiterst beperkt en ook het risico. Toezicht hoefde zich er enkel op te richten vast te stellen of de verwerving en besteding van de middelen rechtmatig hadden plaatsgevonden.

Vanaf midden jaren tachtig is een verandering ingezet die instellingen veel autonomer maakte. Dat is gestart in het hoger onderwijs en vervolgens door de onderwijskolom afgedaald naar uiteindelijk het primair onderwijs, waar het veranderingsproces rond 2006 is afgerond. In de kern kwam deze verandering erop neer dat instellingen bekostigd werden op basis van een beperkt aantal parameters (met aantal leerlingen/studenten als voornaamste) en met behulp van de middelen die op grond daarvan beschikbaar werden gesteld, hun onderwijs volgens een aantal randvoorwaarden moesten inrichten.

Het was daarmee aan de instellingen zelf keuzes te maken voor investeringen en kosten en ervoor te zorgen dat zij die konden dragen. Tussen sectoren bestaan daarbij nog wel verschillen. In principe worden gebouwen in het funderend (primair en voortgezet) onderwijs ter beschikking gesteld door gemeenten, terwijl in mbo en hoger onderwijs de instellingen zelf verantwoordelijk zijn voor de huisvesting. Eigen gebouwen staan daar op de balans en zij lopen daarmee ook meer risico. Ook op het punt van de parameters die gelden voor bekostiging, zitten verschillen. Zo is in het mbo en het hoger onderwijs het aantal gediplomeerden een factor die meetelt.

Het continuïteitstoezicht heeft zich vanaf 2006 ontwikkeld op basis van een aantal modellen om financiële risico's te detecteren. Het eerste model voor het primair en voortgezet onderwijs omvatte drie indicatoren, de liquiditeit (current ratio), rentabiliteit (resultaat/totale baten) en het weerstandsvermogen (eigen vermogen/totale baten, NA/TR volgens de definitie van Tuckman en Chang, 1991) met bijpassende signaleringswaarden (Inspectie van het Onderwijs, 2009). In 2009 is er onderzoek uitgevoerd door de Commissie Vermogensbeheer Onderwijsinstellingen, beter bekend als de Commissie Don (2009). De opdracht aan deze commissie was niet alleen te kijken naar indicatoren voor (continuïteits)risico's bij de instellingen, maar ook naar indicaties van overmatig spaargedrag dat als ondoelmatig middelengebruik wordt beschouwd.

De Commissie Don pleitte voor het laten vervallen van het weerstandsvermogen en het opnemen van de indicatoren solvabiliteit¹⁷ en kapitalisatiefactor naast liquiditeit en rentabiliteit. Voor deze indicatoren zijn zowel onder- als bovengrenzen geformuleerd, gezien het na te streven doel. Dit heeft in 2011 geresulteerd in een nieuw model (Ministerie van OCW, 2011) met daarin deze vier indicatoren. De signaleringsgrens voor de solvabiliteit was hier gelegd op 0,2.

De ondergang, in 2012, van onderwijsinstelling Amarantis (Commissie onderzoek financiële problematiek Amarantis, 2012) heeft vervolgens geleid tot een aantal maatregelen ad hoc en een meer structurele bezinning op het gebruik van de indicatoren op de langere termijn. De maatregelen ad hoc leidden tot een aanpassing van het bestaande model met aanvullende indicatoren en een verscherping van de grenswaarden. Speciaal de grens voor solvabiliteit werd opgehoogd naar 0,3 omdat Amarantis juist scherp rond de eerdere grens opereerde¹⁸. Voorts is een breed onderzoek ingesteld naar de financiële staat van instellingen. Dit heeft geleid tot een forse vergroting van het aantal gedetecteerde en onderzochte instellingen zonder dat goed duidelijk is of daarmee ook meer risico's zijn afgevangen. Recent intern onderzoek (Inspectie van het Onderwijs, 2015) laat zien dat slechts een deel van de indicatoren significante betekenis heeft voor de risicodetectie.

Het model dat resulteerde en dat de inspectie vanaf september 2012 hanteert voor het voortgezet onderwijs, omvat de volgende indicatoren:

- CA/CL (Liquiditeit, current ratio).
- $(NA+VZ)/TA$ (eigen vermogen + voorzieningen/totaal passiva, solvabiliteit 2)¹⁹.
- INTHREE Rentabiliteit (resultaat/totale baten) drie jaar negatief (dichotoom 0 of 1).
- PK/RB (Personele lasten/rijksbijdrage).
- PK/TB (Personele lasten/totale baten).
- Buffer. De buffer is een restwaarde van de kapitalisatiefactor (Commissie Don, 2009)²⁰.

¹⁷ In de praktijk betreft het hier de zogenaamde solvabiliteit-2, berekend als (eigen vermogen + voorzieningen)/totale passiva.

¹⁸ Ondanks het feit dat er slechts sprake is van signaleringsgrenzen, gaan dergelijke waarden, doordat ze door de toezichthouder worden gebruikt, in de praktijk werken als normen voor bijvoorbeeld geldverstrekkers.

¹⁹ Dit wijkt af van de definitie van de overige auteurs, waarbij sprake is van de definitie van solvabiliteit 1 (eigen vermogen/totaal activa) (NA/TA) .

²⁰ De buffer wordt bij Don (2009) berekend uit de zogenaamde kapitalisatiefactor. Die omvat het totale kapitaal minus gebouwen/terreinen. Dit bedrag wordt verminderd met de financieringsfunctie (gelijk

In de dagelijkse praktijk van het toezicht door de inspectie, maakt zij gebruik van grenswaarden bij deze indicatoren ($CA/CL \leq 0,75$, $(NA+VZ)/TA \leq 0,30$, $PK/RB \geq 0,95$, $PK/TB \geq 0,90$, Buffer < 0). In dit hoofdstuk is van die grenswaarden geen gebruik gemaakt.

De meer structurele bezinning kwam voort uit de vaststelling door de commissie dat het toezicht overwegend opereerde op basis van historische gegevens, namelijk jaarrekeningen na inlevering een half jaar na afloop van het boekjaar. Gegevens waren daarmee gemiddeld een jaar oud en minder geschikt om tijdig risico's te signaleren. De gegevens werden pas dan gebruikt voor het detecteren van instellingen die aan een nader onderzoek onderworpen moesten worden. De Commissie Amarantis (2012) betoogde dat het toezicht te retrospectief was ingericht en meer prospectieve elementen zou moeten opnemen. Daarmee bedoelde zij dat zowel financiële als niet-financiële indicatoren zouden moeten worden gezocht met een meer voorspellend karakter, zodat risico's eerder zouden worden gedetecteerd en zodat eerder zou kunnen worden ingegrepen.

Het financiële toezicht door de inspectie leidt ertoe dat instellingen met grote financiële risico's na een diepgaand onderzoek onder AFT worden geplaatst. Dit is een toestand waarbij de instelling nauwkeurig wordt gevolgd tot eventueel de problemen zijn opgelost. Daarmee is de status van AFT een geschikt middel om vast te stellen of de instelling financieel kwetsbaar is. Het kan aldus dienen als alternatief voor andere manieren waarop onderzoekers bepaald hebben dat er sprake is van een financieel kwetsbare organisatie.

3.3 Onderzoeksgegevens

Het databestand bestaat uit de complete financiële gegevens van 1403 schoolbesturen voor primair en voortgezet onderwijs in Nederland van 2006 tot en met 2012. Dit zijn vrijwel alle besturen. Enkele besturen met bijzondere omstandigheden zoals recente oprichting, zijn buiten beschouwing gelaten omdat de gegevens daar nogal afwijken van instellingen die langer bestaan. Instellingen zijn verplicht hun financiële gegevens volgens een vastgesteld format in te dienen in het kader van hun bekostiging. Uit de gegevens worden door de uitvoeringsorganisatie DUO kengetallen berekend, onder meer voor gebruik door de inspectie.

aan de materiële vaste activa met uitzondering van gebouwen/terreinen) en de transactieliquideit (gelijk aan het kort vreemd vermogen). Het restant is de buffer, bedoeld om onzekerheden in de bedrijfsvoering op te kunnen vangen. Om deze bedragen te corrigeren voor de grootte van de instelling, worden ze gedeeld door de totale baten (inclusief rente).

Toen vanaf 2008 de afdeling van het ministerie die voorheen het financiële toezicht uitvoerde, onderdeel werd van de inspectie, is gestart met een systematische analyse van de financiële positie van deze schoolbesturen. Dat heeft erin geresulteerd dat vanaf het jaarverslag 2008 door financiële inspecteurs analyses gemaakt zijn van de financiële situatie van alle schoolbesturen waarbij een of enkele van de kritische indicatoren volgens het inspectiemodel, zoals eerder beschreven, niet aan de signaleringswaarden voldeden. Deze analyses vinden plaats op basis van de jaarverslagen, eventueel aanvullend opgevraagde documenten en gesprekken met het bestuur.

Zoals eerder al vermeld, kan dat vervolgens leiden tot zogenaamd Aangepast Financieel Toezicht (AFT). Het plaatsen onder AFT komt bij een gering aantal instellingen voor. Van de instellingen was, zoals eerder opgemerkt, bijna drie procent in 2014 onder AFT geplaatst. Faillissementen komen vrijwel niet voor²¹. Gordon e.a. (2013) rapporteren in hun onderzoek onder ruim 300.000 non-profitorganisaties een jaarlijks insolventiepercentage van ongeveer 1,6 procent en bij bekostigde onderwijsinstellingen in Nederland is dat nog veel lager. Dit zeer lage aantal maakt feitelijk faillissement geen geschikte afhankelijke variabele, zeker niet met een veel kleinere populatie. AFT biedt die mogelijkheid wel en heeft bovendien het voordeel dat dit niet bepaald is op grond van een enkelvoudige score maar op basis van een uitvoerig boekenonderzoek waarin alle risicoaspecten van de instelling zijn meegenomen. Het bestand met instellingen onder AFT op 1 september 2013 en 1 september 2014 is daarom gebruikt als referentiemaat voor financiële risico's op basis van de jaarrekeningen 2011 en 2012.

3.3.1 Beschrijvende statistiek

Tabel 3.1 laat een overzicht zien van alle variabelen volgens de verschillende methoden en laat zien wat het verschil is tussen de waarden bij de instellingen zonder en met AFT. Daar valt het volgende bij op te merken. Over de administratiekosten (AC/TR) bij Tuckman en Chang (1991) is eerder al opgemerkt dat de definitie ervan niet helder is. Volgens de omschrijving bij de Nederlandse gegevens, omvatten deze administratiekosten een beperkte groep kosten, waaronder ook accountantskosten. Dat zou een verklaring kunnen zijn voor de hogere waarde van de administratiekosten bij de instellingen met AFT (waar accountants het mogelijk drukker hebben). Eerder is ook gesuggereerd

²¹ Ook de eerder genoemde instelling Amarantis is feitelijk niet failliet gegaan, maar is met een hoeveelheid financiële steun opgesplitst en verdergegaan in de vorm van vijf zelfstandige instellingen.

dat er wellicht overheadkosten in het algemeen mee worden bedoeld. Een vergelijking van de gemiddelde overhead (berekend als percentage fte's voor directie en ondersteunend personeel) over 2009-2013 levert op dat dit 23,7 procent is bij scholen zonder AFT en ruim 24 procent bij scholen met AFT. Dit verschil is niet significant. Het getal van rond 24 procent is meer in lijn met het percentage in Tuckman en Chang (1991). Dit overheadpercentage is evenwel niet voor alle instellingen bekend, zodat toch van de administratiekosten gebruik wordt gemaakt.

Tabel 3.1: Beschrijvende statistiek op basis van jaarrekening 2011

	N	Minimum	Maximum	Gemiddeld	Std. deviatie	AFT=0 (N 1371)	AFT=1 (N 32)
Tuckman en Chang							
NA/TR	1403	-,701000	2,238000	,34456593	,230413269	,35198614**	,02665625
RCI	1403	,418766	,998019	,85069187	,086811187	,85147352*	,81720301
AC/TR	1403	,000947	,204805	,03056137	,016907067	,03042077*	,03658522
NI/TR	1403	-,393000	,489000	-,00588168	,048975483	-,00478191**	0,053
Altman							
WC/TA	1403	-2,256955	,930271	,31701631	,286082671	,33032015**	-,25297019
NA/TA	1403	-3,300000	,930000	,57978617	,235259798	,59366156**	-,0146875
EBIT/TA	1403	-1,565197	,571956	-,02733098	,107489701	-,02316892**	-,20564921
TR/TA	1403	,278057	11,154778	2,14560178	,948307483	2,10578778**	3,85138298
Ohlson							
SIZE	1403	10,864071	19,432140	14,68223048	1,321586401	14,69299431*	14,22106728
TL/TA	1403	,070000	4,300000	,42021383	,235259798	,40633844**	1,0146875
WC/TA	1403	-2,256955	,930271	,31701631	,286082671	,33032015**	-,25297019
CL/CA	1403	-,724638	16,340040	,62519732	,882850586	,5940903**	1,95793863
OENEG	1403	0	1	,01	,109	,01**	0,28
NI/TA	1403	-1,593415	,567972	-,01824782	,107515341	-,01380891**	-,20842734
FFO/TL	1403	,389600	15,186638	6,00351064	2,045704724	6,03550845**	4,63260446
INTWO	1403	0	1	,41	,492	,4**	,66
CHIN	1403	-1,000000	1,000000	,07522263	,706510098	,07780723	-,03551116
Inspectie							
INTHREE	1403	0	1	,23	,420	,22**	,47
NA+VZ/TA	1403	-3,215583	,950000	,68490146	,218840697	,69893706**	,0835639
CA/CL	1403	-1,380000	18,250000	2,88868618	2,209579200	2,93727989**	,80674913
BUFFER	1403	-,413000	1,364324	,24976532	,213856806	,25611651**	-,03860795
PK/RB	1403	,400000	1,499000	,89976232	,086562848	,89875624**	,94286634
PK/TB	1403	,396349	1,134026	,82073744	,054841218	,82038117	,83600153

Verskil 0/1 AFT ** significant 0,01, * significant 0,05 (ANOVA, dichotome variabelen Mann-Whitney)

3.3.2 Wijze van analyseren van de gegevens

Voor de analyse is primair gebruik gemaakt van logistische regressie. Daarbij zijn de onafhankelijke variabelen uit de jaarrekeningen van twee verschillende uitgangsjaren gekozen, namelijk 2010 en 2011. Met deze cijfers zijn afzonderlijk regressies gemaakt naar de AFT's op grond van de jaarrekening 2011. Dit levert twee regressievergelijkingen op. Vervolgens is beschouwd of en in welke mate de invulling van deze regressievergelijkingen met de cijfers uit 2011 een adequate voorspelling oplevert van de AFT's in 2012. Tabel 3.2 laat dit schematisch zien.

Tabel 3.2: Stappen in de analyse

Jaarrekening 2010	Jaarrekening 2011
Regressie AFT2011 op variabelen JR 2010	Regressie AFT2011 op variabelen JR 2011
Invulling gevonden regressievergelijking in JR 2011 ter voorspelling AFT's 2012	Invulling gevonden regressievergelijking in JR 2011 ter voorspelling AFT's 2012
Vergelijking met feitelijke AFT's 2012	Vergelijking met feitelijke AFT's 2012

Met deze werkwijze vervalt de noodzaak om de populatie op te delen in een onderzoeksgroep en een controlegroep. Als controlegroep wordt immers dezelfde groep gebruikt, maar dan een jaar later. Gezien het continue karakter van het financiële toezicht, is dat een werkwijze die daar goed bij aansluit. De reden om de vergelijking zowel vanuit jaarrekening 2010 als jaarrekening 2011 op te stellen, is gelegen in de volgende overweging. De uiteindelijke toepassing van het model is erop gebaseerd dat vanuit het ene jaar een voorspelling wordt geformuleerd voor het volgende jaar. Bij de vergelijking op grond van jaarrekening 2010 wordt een dergelijke benadering van meet af aan gebruikt. Deze wordt vervolgens een jaar later toegepast om 2012 vanuit 2011 te voorspellen. Het is denkbaar dat dit tijdsaspect van invloed is op de kwaliteit van die voorspelling.

Dat toekomstelement zit er niet in als de regressie primair wordt gebaseerd op de jaarrekening 2011 en vervolgens, op grond van diezelfde jaarrekening geprojecteerd wordt op een volgend jaar. Daartegenover staat het eerder al door Duda en Schmidt (2010) gemelde verschijnsel dat de financiële indicatoren kort voor een faillissement meer "noisy" zijn, waardoor een vergelijking op grond van de meer recente gegevens uit

2011 toch een betere voorspelling zou kunnen opleveren. Door volgens beide methoden te werken, kan dat mogelijk inzichtelijk worden gemaakt.

Deze analyses worden uitgevoerd met de onderscheidingswaarden 0,1 en 0,03 om na te gaan wat de effecten daarvan zijn op de kwaliteit van de voorspelling. De 0,1 is gebaseerd op de bevindingen van Greenlee en Trussel (2000). De 0,03 ligt dicht bij het percentage AFT's over de jaarrekening 2012 (2,7 procent). Tenslotte wordt gekeken naar de meerwaarde die een hiërarchisch besluitvormingsmodel kan hebben. Dit gebeurt door de uitkomsten van telkens twee modellen aan elkaar te verbinden door middel van een logische of-vergelijking: als een instelling in de ene of in de andere als risicovol wordt omschreven, wordt zij in de gecombineerde lijst als risicovol beschouwd.

De beoordeling van de uitkomsten van de modellen en combinaties vindt plaats door de bepaling van de accuraatheid van het model. Daarbij gelden twee belangrijke overwegingen. De eerste is de effectiviteit. Daarmee wordt bedoeld dat er minimaal type-2-fouten worden gemaakt. De tweede is die van de efficiency, waarbij het erom gaat zo min mogelijk instellingen ten onrechte van een nader onderzoek te voorzien, dus type-1-fouten te maken²². Dit leidt tot de volgende definitie:

- Efficiency = $(1 - \text{aantal ten onrechte te onderzoeken objecten} / \text{totaal aantal objecten}) * 100$ (procent)
- Effectiviteit = $(\text{aantal als risicovol geïdentificeerde objecten} / \text{totaal aantal risicovolle objecten}) * 100$ (procent)

In het ideale geval is de efficiency 100 procent. Er worden geen objecten ten onrechte onderzocht. Iets vergelijkbaars geldt voor de effectiviteit. Alle risicovolle objecten worden als zodanig geïdentificeerd. Om de modellen goed te kunnen vergelijken, worden beide verenigd in een getal, de accuraatheid, door ze met elkaar te vermenigvuldigen. In het ideale geval is ook de accuraatheid 100 procent. Door deze werkwijze is de relatieve invloed van type-2-fouten in het totaal van de beoordeling veel zwaarder dan van type-1-fouten. Dit weerspiegelt de veel grotere kosten die er voor de toezichthouder zijn verbonden aan fouten van type-2. Ter controle zijn de uitkomsten van deze wijze van beoordelen achteraf vergeleken met de areas under the curve uit een ROC-analyse (Metz, 1978). De correlatie hiertussen was 0,998.

²² Deze begrippen lijken sterk op de begrippen specificiteit en sensitiviteit. Effectiviteit is gelijk aan specificiteit. Efficiency wijkt echter af. Het gaat hierbij om alle gevallen waarin ten onrechte een verdergaand onderzoek is uitgevoerd. Deze score is in het algemeen wat lager dan die voor sensitiviteit.

3.4 Onderzoeksbevindingen

3.4.1 Resultaten van de regressies

In de tabellen 3.3 en 3.4 zijn de resultaten van de regressies van 2011 op 2011, respectievelijk 2010 op 2011 met de verschillende modellen opgenomen. Bij de Ohlson-variabelen passen daarbij de volgende opmerkingen. De toepassing van het volledige Ohlson-model op de gegevens leidde tot problemen met multicollineariteit. Op grond van Ohlson (1980) en Duda en Schmidt (2010) is daarom het aantal indicatoren gereduceerd tot die welke als statistisch significant zijn te beschouwen. Met deze vier indicatoren deed het probleem zich niet meer voor.

Omdat er sprake is van instellingen uit verschillende onderwijssectoren en met een verschillende grootte, is met behulp van twee controlevariabelen nagegaan of dit van invloed is op de mate waarin instellingen risico's lopen op AFT. Daarbij bleek dat noch de sector, noch de omvang van de instelling daarop een significante invloed had. Met betrekking tot de grootte komt dit overeen met de bevinding op basis van het model van Ohlson, waarin de variabele SIZE ook geen significante bijdrage levert.

Tabel 3.3: Regressie AFT-2011 op basis van jaarrekening 2011

Omschrijving	Verwacht teken	Tuckman en Chang	Altman	Ohlson	Inspectie
Intercept		-0,344	-3,026**	-5,596*	-7,597*
NA/TR	-	-14,600**			
RCI	+	-1,321			
AC/TR	-	-4,720			
NI/TR	-	-14,980**			
WC/TA	-		-2,467**	-2,570**	
NA/TA	-		-3,283**		
EBIT/TA	-		-4,821**		
TR/TA	-		0,198		
SIZE	-			-0,021	
TL/TA	+			3,499**	
NI/TA	-			-5,499**	
INTHREE	+				-0,579
(NA+VZ)/TA	-				-4,096**
CA/CL	-				-0,571
BUFFER	-				-4,482
PK/RB	+				1,220
PK/TB	+				7,737
-2 log likelihood		182,759	182,122	182,162	175,085
Nagelkerke R ²		0,428	0,430	0,430	0,453
Accuraatheid		0,60	0,58	0,61	0,67

** P<0.01, * P<0.05, accuraatheid bij onderscheidingswaarde 0,1

Tabel 3.4: Regressie AFT-2011 op basis van jaarrekening 2010

Omschrijving	Verwacht teken	Tuckman en Chang	Altman	Ohlson	Inspectie
Intercept		0,669	-3,093**	-5,748*	-9,293**
NA/TR	-	-9,619**			
RCI	+	-2,776			
AC/TR	-	-8,308			
NI/TR	-	-7,831*			
WC/TA	-		-0,147	-0,433	
NA/TA	-		-2,994**		
EBIT/TA	-		-1,239		
TR/TA	-		0,264		
SIZE	-			0,013	
TL/TA	+			3,317**	
NI/TA	-			-1,956*	
INTHREE	+				-0,390
(NA+VZ)/TA	-				-3,306**
CA/CL	-				-0,112
BUFFER	-				-2,082
PK/RB	+				3,206
PK/TB	+				6,387
-2 log likelihood		227,456	241,594	244,345	223,683
Nagelkerke R ²		0,275	0,226	0,217	0,268
Accuraatheid		0,46	0,34	0,34	0,44

** P<0.01, * P<0.05, accuraatheid bij onderscheidingswaarde 0,1

3.4.2 Voorspelling van de resultaten op basis van jaarrekening 2012

De coëfficiënten die verkregen zijn uit de hierboven uitgevoerde regressies, zijn verwerkt in vergelijkingen waarmee de kansen berekend zijn op AFT op grond van de jaarrekening 2012. De algemene vergelijking hiervan ziet eruit als volgt:

$P = 1/(1 + e^{-z})$, waarbij z de uitkomst is van de vergelijking met de ingevulde coëfficiënten.

De regressievergelijkingen op grond van jaarrekening 2011 zijn per model de volgende:

Tuckman en Chang: $z = -0,344 - 14,6 \cdot \text{NA/TR} - 1,321 \cdot \text{RCI} - 4,72 \cdot \text{AC/TR} - 14,98 \cdot \text{NI/TR}$

Altman: $z = -3,026 - 2,467 \cdot WC/TA - 3,283 \cdot NA/TA - 4,821 \cdot EBIT/TA + 0,198 \cdot TR/TA$

Ohlson: $z = -5,596 - 2,570 \cdot WC/TA - 0,021 \cdot SIZE + 3,499 \cdot TL/TA - 5,499 \cdot NI/TA$

Inspectie: $z = -7,597 - 0,579 \cdot INTHREE - 4,096 \cdot (NA+VZ)/TA - 0,571 \cdot CA/CL - 4,482 \cdot BUFFER + 1,220 \cdot PK/RB + 7,737 \cdot PK/TB$

In tabel 3.5 staan de voorspellingen van AFT-2012 met behulp van deze regressievergelijkingen met onderscheidingswaarde 0,1.

Tabel 3.5: Voorspelling 2012 op basis van regressie 2011 en jaarrekening 2011 (onderscheidingswaarde 0,1)

Methode	Efficiency	Effectiviteit	Accuraatheid
Tuckman en Chang	0,97	0,61	0,59
Altman	0,98	0,55	0,54
Ohlson	0,98	0,58	0,57
Inspectie	0,97	0,63	0,61
Gemiddeld	0,975	0,593	0,578

In tabel 3.6 zijn de voorspellingen opgenomen op grond van de jaarrekening 2010. Hiervoor zijn de volgende vergelijkingen gebruikt:

Tuckman en Chang: $z = -0,669 - 9,619 \cdot NA/TR - 2,776 \cdot RCI - 8,308 \cdot AC/TR - 7,831 \cdot NI/TR$

Altman: $z = -3,093 - 0,147 \cdot WC/TA - 2,994 \cdot NA/TA - 1,239 \cdot EBIT/TA + 0,264 \cdot TR/TA$

Ohlson: $z = -5,748 - 0,433 \cdot WC/TA - 0,013 \cdot SIZE + 3,317 \cdot TL/TA - 1,956 \cdot NI/TA$

Inspectie: $z = -9,293 - 0,390 \cdot INTHREE - 3,306 \cdot (NA+VZ)/TA - 0,112 \cdot CA/CL - 2,082 \cdot BUFFER + 3,206 \cdot PK/RB + 6,387 \cdot PK/TB$

Dat laat in het algemeen het beeld zien van een lagere graad van accuraatheid dan in tabel 3.5.

Tabel 3.6: Voorspelling 2012 op basis van regressie 2010 en jaarrekening 2011 (onderscheidingswaarde 0,1)

Methode	Efficiency	Effectiviteit	Accuraatheid
Tuckman en Chang	0,98	0,58	0,57
Altman	0,99	0,45	0,44
Ohlson	0,99	0,39	0,39
Inspectie	0,98	0,63	0,62
Gemiddeld	0,985	0,513	0,505

3.4.3 **Aanpassing van de onderscheidingswaarde**

De bovenstaande uitkomsten zijn bepaald op grond van een onderscheidingswaarde van 0,1 voor het onderscheid tussen wel of niet als risicovol classificeren van de instellingen. Zoals eerder als is opgemerkt, is de keuze van een onderscheidingswaarde van groot belang voor de verdeling van fouten de eerste, respectievelijk de tweede soort. Om na te gaan of dit een grotere accuraatheid oplevert, is de voorspelling voor 2012 nogmaals gedaan met de waarde 0,03 die overeenkomt met de proportie risicovolle objecten volgens jaarrekening 2012. De resultaten daarvan staan hieronder in de tabellen 3.7 en 3.8.

Tabel 3.7: Voorspelling 2012 op basis van regressie 2011 en jaarrekening 2011 (onderscheidingswaarde 0,03)

Methode	Efficiency	Effectiviteit	Accuraatheid
Tuckman en Chang	0,88	0,82	0,72
Altman	0,91	0,79	0,72
Ohlson	0,91	0,79	0,72
Inspectie	0,87	0,84	0,73
Gemiddeld	0,893	0,810	0,723

Tabel 3.8: Voorspelling 2012 op basis van regressie 2010 en jaarrekening 2010 (onderscheidingswaarde 0,03)

Methode	Efficiency	Effectiviteit	Accuraatheid
Tuckman en Chang	0,83	0,84	0,70
Altman	0,89	0,79	0,70
Ohlson	0,88	0,82	0,72
Inspectie	0,81	0,87	0,71
Gemiddeld	0,853	0,830	0,708

Duidelijk is dat, zoals te verwachten was, het aantal fouten van de eerste soort door deze aanpassing aanzienlijk toeneemt. De efficiency-scores weerspiegelen dat treffend. Deze zijn ongeveer 0,1 lager dan bij de onderscheidingswaarde van 0,1. Als echter gekeken wordt naar de maat van de accuraatheid, waarin de fouten van de tweede soort afzonderlijk zijn verdisconteerd, levert dat een heel ander beeld. Deze neemt gemiddeld op grond van de regressie uit 2011 toe van 0,578 naar 0,723 en op grond van de regressie uit 2010 van 0,505 naar 0,708.

Als het gaat om de waarde van de verschillende methoden, zijn de verschillen daartussen niet groot. Opvallend is vooral de wat lagere score

van de methoden van Altman en Ohlson in tabel 5. Dat is daar een gevolg van een duidelijk lagere effectiviteit.

3.4.4 Gecombineerde besluitvormingsmethoden

Behalve door het aanscherpen van de technische methode, is het ook denkbaar de accuraatheid van de voorspelling te vergroten door aanpassing van het besluitvormingsproces. Zoals ook eerder al is besproken onder verwijzing naar Deng en Hendrikse (2014), is de gangbare beslissing over de vraag of er wel of niet sprake is van een risicogeval er een die in een stap wordt genomen. Door twee detectiemethoden naast (of na) elkaar te plaatsen met een verschillende beoordelingsoptiek, en die zelfstandig te laten besluiten, zou een verdere verbetering mogelijk zijn. De afzonderlijke beoordelingen worden vervolgens gecombineerd, in dit geval door elke instelling als risicogeval te bestempelen die door een van beide methoden als zodanig wordt aangemerkt. In de tabellen 3.9 en 3.10 is het resultaat daarvan weergegeven. Tabel 3.9 geeft de uitkomsten weer met de onderscheidingswaarde 0,1. Linksboven staan combinaties van methoden met de regressievergelijkingen op basis van 2011. Rechtsonder zijn dat de uitkomsten met regressievergelijkingen op basis van 2010. In tabel 3.10 is hetzelfde weergegeven, maar dan voor de onderscheidingswaarde 0,03.

Tabel 3.9: Uitkomsten van combinaties onderscheidingswaarde 0,1

2011-2011		Inspectie	Ohlson	Altman	T&C
T&C	efficiency	0,96	0,96	0,96	
	effectiviteit	0,71	0,68	0,66	
	accuraatheid	0,68	0,66	0,63	
Altman	efficiency	0,96	0,98		0,98
	effectiviteit	0,68	0,58		0,58
	accuraatheid	0,66	0,56		0,56
Ohlson	efficiency	0,96		0,99	0,98
	effectiviteit	0,68		0,47	0,58
	accuraatheid	0,66		0,47	0,57
Insp.	efficiency		0,97	0,97	0,97
	effectiviteit		0,63	0,66	0,68
	accuraatheid		0,62	0,64	0,66
				2010-2011	

Tabel 3.10: Uitkomsten van combinaties onderscheidingswaarde 0,03

2011-2011		Inspectie	Ohlson	Altman	T&C
T&C	efficiency	0,84	0,87	0,87	
	effectiviteit	0,89	0,84	0,84	
	accuraatheid	0,75	0,73	0,73	
Altman	efficiency	0,85	0,91		0,83
	effectiviteit	0,87	0,79		0,84
	accuraatheid	0,74	0,71		0,69
Ohlson	efficiency	0,86		0,86	0,80
	effectiviteit	0,87		0,84	0,89
	accuraatheid	0,74		0,73	0,72
Insp.	efficiency		0,79	0,79	0,76
	effectiviteit		0,89	0,89	0,92
	accuraatheid		0,70	0,71	0,70
				2010-2011	

3.4.5 Resultaten van de verschillende methoden

De bovenstaande analyses leverden een groot aantal uitkomsten op (N=40) gebaseerde op verschillende invalshoeken. Om uitspraken te kunnen doen over resultaten van de modellen en de gebruikte methoden, zijn deze met de uitkomsten (de scores op accuraatheid) in een tabel geplaatst, en is via regressie beoordeeld in welke mate de methoden bijdragen aan de uitkomst.

In tabel 3.11 zijn de resultaten te zien van die regressies met als verklarende variabelen de onderscheidingswaarde (0,1 of 0,03), uitgangsjaar (2010 of 2011) en enkelvoudig/hiërarchisch. De tabellen laten zien dat de belangrijkste, verklarende variabele in de onderscheidingswaarde zit ($p < 0.000$). De beide andere variabelen leveren eveneens significante bijdragen, zij het op een iets lager niveau (respectievelijk 0,018 en 0,019).

Tabel 3.11: Vergelijking modellen met verschillende variabelen

Variable	Onder-scheiding (=model 1)	Model 1 met uit-gangsjaar (=model 2)	Model 2 met enkel/meer-voudig
Intercept	0,585 (0,013)	0,564 (0,014)	0,541 (0,017)
Onderscheiding	0,134 (0,018)	0,134 (0,017)	0,134 (0,016)
Uitgangsjaar		0,042 (0,017)	0,042 (0,016)
Enkel/meer-voudig			0,039 (0,16)
σ^2	0,003	0,003	0,002
R^2	0,596	0,654	0,703
N	40	40	40

De verschillen tussen de beide onderscheidingswaarden zijn in tabel 3.12 per indicator inzichtelijk gemaakt. Zoals te verwachten is, neemt de efficiency door deze keuze af, maar de effectiviteit neemt zodanig toe dat dat resulteert in een veel hogere accuraatheid. Alle verschillen zijn significant (Mann-Whitney, $p < 0,01$). Deelvraag drie is daarmee van een antwoord voorzien.

Tabel 3.12: Verschillen tussen de onderscheidingswaarden

		Efficiency	Effectiviteit	Accuraatheid
0,1 N=20	Gemiddelde	,9730	,6005	,5850
	Std. Deviatie	,01081	,08451	,07810
0,03 N=20	Gemiddelde	,8505	,8465	,7185
	Std. Deviatie	,04347	,03911	,01599
Totaal N=40	Gemiddelde	,9118	,7235	,6517
	Std. Deviatie	,06946	,14050	,08756

Soortgelijke verschillen zijn te zien tussen de waarden van de variabelen op basis van de uitgangsjaren (tabel 3.13). Anders dan bij de onderscheidingswaarde is hier niet alleen de efficiency maar ook de effectiviteit hoger. De verschillen zijn echter niet dusdanig groot dat ze tot significantie leiden. Deelvraag twee wordt daarmee niet eenduidig beantwoord.

Tabel 3.13: Verschillen tussen de uitgangsjaren

		Efficiency	Effectiviteit	Accuraatheid
2010	Gemiddelde	,9020	,7120	,6310
N=20	Std. Deviatie	,08532	,16710	,10125
2011	Gemiddelde	,9215	,7350	,6725
N=20	Std. Deviatie	,04923	,11100	,06766
Totaal	Gemiddelde	,9118	,7235	,6517
N=40	Std. Deviatie	,06946	,14050	,08756

Detailresultaten van het verschil tussen een enkelvoudige of een samengestelde beoordelingsmethode worden weergegeven in tabel 3.14. De gemiddelde waarden komen overeen met de verwachtingen. Deze verschillen zijn echter over het geheel genomen, niet significant. Dat wordt voor een belangrijk deel veroorzaakt door de invloed van de onderscheidingswaarden.

Tabel 3.14: Verschillen enkelvoudige/meervoudige besluitvorming

		Efficiency	Effectiviteit	Accuraatheid
Enkelv.	Gemiddelde	,9263	,6863	,6281
N=16	Std. Deviatie	,06087	,15187	,10635
Meerv.	Gemiddelde	,9021	,7483	,6675
N=24	Std. Deviatie	,07431	,12970	,07054
Totaal	Gemiddelde	,9118	,7235	,6517
N=40	Std. Deviatie	,06946	,14050	,08756

Als afzonderlijk voor elke onderscheidingswaarde wordt gekeken, zijn de verschillen bij een onderscheidingswaarde van 0,1 voor alle maten significant ($p < 0,05$). Bij een onderscheidingswaarde van 0,03 is het beeld echter wisselend. De efficiency neemt significant af ($p < 0,05$) en de effectiviteit neemt significant toe ($p = 0,01$) maar deze veranderingen compenseren elkaar dusdanig dat er geen significante toename van de accuraatheid meer is. Het lijkt erop dat met beide effecten een plafond van accuraatheid wordt bereikt. Deelvraag vier is hiermee gedeeltelijk positief beantwoord.

Interessant is verder of een combinatie van ongelijksoortige beslissingseenheden een beter resultaat haalt dan van gelijksoortige beslissingseenheden. Eerder is opgemerkt dat de modellen van Altman en Ohlson voor bedrijven zijn ontwikkeld en meer op de kapitaalkant van de balans zijn georiënteerd, terwijl de modellen van Tuckman en Chang en de Inspectie voor non-profitorganisaties zijn ontwikkeld en meer naar de batenkant van de resultatenrekening kijken. Analyse laat zien dat een

combinatie van ongelijksoortige beslissingseenheden weliswaar een iets grotere effectiviteit en accuraatheid laat zien (bij vrijwel gelijkblijvende efficiency), maar dat geen van deze verschillen significant is. Deelvraag vijf wordt daarmee niet eenduidig beantwoord.

Hetzelfde geldt voor het onderscheid naar deze soorten modellen. Tabel 3.15 laat de verschillen tussen de soorten zien. De modellen van Altman/Ohlson laten een lagere accuraatheid zien dan de beide andere, maar door het lage aantal waarnemingen is geen van de verschillen significant. Hier is er dus geen eenduidig antwoord op deelvraag een.

Tabel 3.15: Verschillen in de aard van de ratiomodellen

		Efficiency	Effectiviteit	Accuraatheid
T&C/Inspectie N=12	Gemiddelde	,9017	,7517	,6700
	Std. Deviatie	,07895	,12350	,05924
Altman/Ohlson N=12	Gemiddelde	,9375	,6533	,6058
	Std. Deviatie	,04883	,16599	,12595
Totaal N=24	Gemiddelde	,9196	,7025	,6379
	Std. Deviatie	,06676	,15164	,10168

3.4.6 Analyse van de type-2-fouten

Ten einde meer inzicht te krijgen in de verschillen tussen de resultaten, is een inhoudelijke analyse uitgevoerd op de uitkomsten van de combinaties in tabel 3.10. Onderzocht is welke besturen wel aangepast financieel toezicht hadden ontvangen maar door een of meer van die combinaties niet werden gedetecteerd, dus een type-2-fout opleverden. Dit leverde een lijst van negen besturen op. Voor de diepere, inhoudelijke analyse van deze besturen is gebruik gemaakt van de (openbare) inspectierapporten (Inspectie van het Onderwijs, 2016) die over deze besturen zijn gemaakt.

Drie besturen scoorden op geen enkele van de onderzochte combinaties van modellen onvoldoende. Het eerste van deze drie maakt deel uit van een groter conglomeraat van besturen. Het overkoepelend deel, waarin deze instelling is geconsolideerd, is onder aangepast financieel toezicht geplaatst, en daarmee is dat ook gaan gelden voor de besturen die eronder vallen, hoewel hun eigen jaarrekeningen daar geen aanleiding toe geven. Het is voor dit bestuur dus terecht dat het niet is gedetecteerd.

Het tweede bestuur kent een meerjarige, negatieve rentabiliteit. Dit is een indicator in het model van de Inspectie (INTHREE, driejarig negatief). In de regressievergelijkingen met dit model heeft deze rentabiliteit een lage

coëfficiënt in vergelijking met de andere indicatoren. Aangezien de instelling voor de veel zwaarder tellende solvabiliteit en de buffer nog positief scoorde, is deze instelling buiten beschouwing gebleven.²³

Het derde bestuur kende in 2011, het jaar waarop de analyse is gebaseerd, enkel positieve kengetallen. Het gaat daarbij echter om een nieuw schoolbestuur in opbouw. Deze besturen vertonen vaak sterk schommelende resultaten, zeker als ze groeien, terwijl ze in het algemeen nog een vrij beperkte balans kennen. Een dergelijke groei kan, in combinatie met een zogenaamde T-1-bekostiging (bekostiging op grond van het leerlingenaantal van een jaar eerder) leiden tot vrij ingrijpende ontwikkelingen op de balans, onder meer door de noodzaak personeel een bepaalde tijd voor te financieren. In het geval van dit bestuur leidde dat tot een terugloop van de solvabiliteit van 0,33 in 2011 naar 0,08 in 2012. Een dergelijke situatie kan een financieel model onmogelijk voorzien. Wel kan door verbreding van het model naar niet-financiële indicatoren, zoals in dit geval de leeftijd van het bestuur, de toezichthouder actiever op de risico's worden geattendeerd.

Bij de zes overige besturen was sprake van aangepast financieel toezicht en werd dat niet gesignaleerd door een of meer van de combinaties van modellen. Bij de eerste drie was er slechts een model dat ze detecteerde. Het eerste bestuur werd alleen door het inspectiemodel gesignaleerd met de hoge onderscheidingswaarde (0,03). Analyse van het inspectierapport laat zien dat dit veroorzaakt is door een negatieve buffer. Deze telt met een vrij zware coëfficiënt mee in de regressievergelijking. Het verschijnsel dat deze met de lagere onderscheidingswaarde (0,1) niet is gesignaleerd, laat zien dat het om een marginaal effect gaat. Dat wordt bevestigd door het gegeven dat het aangepast financieel toezicht van dit bestuur een jaar later al weer is opgeheven.

Twee andere instellingen werden alleen door het model van Tuckman en Chang gesignaleerd en uiteraard in de combinaties daarvan met de andere. Beide besturen hebben te maken (gehad) met bestuurlijke problemen die zich over het hele functioneren van de instelling uitstrekten en die mede reden zijn geweest de instellingen ook onder aangepast financieel toezicht te plaatsen. Dat verklaart waarom ze niet gesignaleerd zijn door het inspectiemodel maar toch onder AFT zijn geplaatst. Uiteraard zijn bestuurlijke problemen geen onderdeel van het model van Tuckman en Chang. Omdat het maar om twee instellingen gaat, is het niet eenvoudig te bepalen waar dat door wordt veroorzaakt. De

²³ Merk hierbij op dat de inspectie in haar reguliere werk niet gebruik maakt van logistische regressie, maar elke variabele toetst aan een signaleringswaarde, waardoor deze instelling wel werd gevonden.

doorslaggevende reden hier lijkt te zijn de relatief lage waarde van NA/TR, een indicator die ook wel wordt aangeduid als het weerstandsvermogen. Deze telt met een zware coëfficiënt mee in dit model.

Dit vermoeden wordt bevestigd door een volgend bestuur van de laatste groep van drie die door meer dan een model werden gevonden. Dit eerste bestuur werd juist alleen door het model van Tuckman en Chang niet gevonden en door alle andere wel. Bij dit bestuur was sprake van een lage liquiditeit, die in enigerlei vorm in alle andere modellen zit (bij Altman in de vorm van werkkapitaal/totale activa) maar bij Tuckman en Chang ontbreekt. Het weerstandsvermogen, een zwaar meetellende indicator bij Tuckman en Chang, was juist wel voldoende.

Het voorlaatste bestuur werd gedetecteerd door de modellen van Ohlson en de inspectie. Dit bestuur is, zoals ook in een eerder geval, een kort tevoren gestart bestuur met nog vrij instabiele indicatoren. In 2011, het eerste jaar van de jaarrekening, waren het weerstandsvermogen en rentabiliteit positief, waardoor het model van Tuckman en Chang het niet signaleerde. Daarentegen was juist de rentabiliteit in dat jaar zeer hoog, een indicator die met een zware coëfficiënt in het model van Altman meetelt.

Bij het laatste bestuur was sprake van een bijzondere situatie, aangezien het geconfronteerd werd met vorderingen uit de overname van een ander bestuur dat eerder al failliet was gegaan. Dit leidde vervolgens tot personeelsclaims en het weigeren van uitkeringen waardoor het bestuur ervoor moest opdraaien (via de post personeelslasten). Dit heeft geleid tot detectie door het inspectiemodel in verband met de hoogte daarvan (PK/TB). In Tuckman en Chang speelt de negatieve rentabiliteit (NI/TR) een zware rol. Hoewel deze indicatoren in een iets andere vorm ook meewegen in de modellen van Altman en Ohlson, doen zij dat minder zwaar. Om die reden detecteerden alleen de modellen van Tuckman en Chang en de inspectie dit bestuur, en alleen met de hoge onderscheidingswaarde van 0,03.

3.5 Conclusie en discussie

Het onderzoek draaide om de vraag of de financiële toezichthouder met behulp van financiële ratiomodellen eerder in staat is financiële risico's te detecteren. Deze vraag is verder toegespitst op een aantal condities die daarbij aan de orde zijn, zoals een zo gering mogelijk aantal fouten van de tweede soort als een vrij harde ontwerpvoorwaarde, maar gelijktijdig uit overwegingen van efficiency niet teveel fouten van de eerste soort.

Bekeken is hoe deze verhouding het best in beeld kon worden gebracht en vooral, welke ontwerpvoorwaarden tot een optimaal resultaat zouden leiden. Daarbij is gebruik gemaakt van een rekenwijze die onderscheid maakt in efficiency en effectiviteit van de methode en die met elkaar in verband brengt in de vorm van accuraatheid. Dit leidt ertoe dat veel preciezer de afweging tussen fouten van de eerste en tweede soort kan worden gemaakt.

De keuze om de AFT's van 2011 ook te regresseren op de financiële gegevens van 2010 was ingegeven door de overweging dat het tijdverloop een relevant gegeven zou kunnen zijn en dat dat van belang zou zijn bij de voorspelling van 2012 op basis van de financiële gegevens van 2011. Die veronderstelling wordt door de resultaten niet bevestigd. De regressie op 2011 blijkt in alle opzichten een betere voorspelling op te leveren, maar het verschil tussen beide methoden is niet significant.

Het bepalen van een geschikte onderscheidingswaarde is een belangrijk gegeven, met name waar het beperken van een bepaald type fout een belangrijke voorwaarde is. De vaak standaard gebruikte onderscheidingswaarde van 0,5 is, gezien de verdeling van de kanswaarden en de doelstelling om het aantal fouten van de tweede soort te minimaliseren, in elk geval ongeschikt. Een onderscheidingswaarde van 0,1 maakt dat onderscheid al beter. Een onderscheidingswaarde van 0,03 die overeenkomt met de proportie risicogeveallen in de populatie blijkt dit nog duidelijk beter te doen. Weliswaar neemt het aantal fouten van de eerste soort er behoorlijk door toe, maar dat wordt meer dan gecompenseerd door de correcte detectie van risicogeveallen. De totale accuraatheid neemt er door toe.

De combinatiemethode laat zien dat door een hiërarchische besluitvorming de prestaties van de methode verder kunnen worden verbeterd. Wel blijkt dat het effect ervan afneemt als een zeker plafond wordt bereikt. Op de vraag of de oriëntatie van de modellen van invloed is, alsmede op de vraag of een combinatie van ongelijksoortige oriëntaties leidt, heeft de analyse geen eenduidige antwoorden opgeleverd.

De werkwijzen laten wel zien dat het toepassen van dergelijke modellen door een financiële toezichthouder als de Inspectie van het Onderwijs, zinvol kan zijn. Als het mogelijk is om (in het gunstigste geval) rond 90 procent van de instellingen met risico's een jaar eerder te detecteren, kan eerder worden ingegrepen en is de kans op verder afglijden van de instelling geringer.

De analyse van de type-2-fouten laat zien dat de gronden waarop modellen afwijkende resultaten geven, divers zijn en het geringe aantal maakt het lastig daaruit een algemene conclusie te trekken. Opvallend is echter dat het model van Tuckman en Chang enkele vondsten doet die in geen van de andere modellen terugkomen, dankzij de indicator weerstandsvermogen (NA/TR). Dit model is als enige overwegend georiënteerd op de opbrengsten (TR) in plaats van op de activa (TA). Hoewel het model voor de Nederlandse situatie wat opmerkelijke bijzonderheden heeft, lijkt die primaire oriëntatie een zinvol aangrijpingspunt voor verder onderzoek naar modellen voor bekostigde non-profitorganisaties. Voor dergelijke instellingen vormt immers de vrij zekere stroom aan inkomsten een belangrijker basis voor hun voortbestaan dan de activa.

Literatuur

- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589-609.
- Chabotar, K. J. (1989). Financial ratio analysis comes to nonprofits. *The Journal of Higher Education*, 60(2), 188-208.
- Commissie onderzoek financiële problematiek Amarantis. (2012). *Autonomie verplicht. rapport onderzoek financiële problematiek Amarantis*. Den Haag:
- Commissie Vermogensbeheer Onderwijsinstellingen (Commissie Don). (2009). *Financieel beleid van onderwijsinstellingen. Rapport van de commissie vermogensbeheer onderwijsinstellingen*. Den Haag: Ministerie van OCW.
- Deng, W., & Hendrikse, G. W. J. (2014). *Managerial vision bias and cooperative governance*. Ongepubliceerd manuscript.
- Duda, M., & Schmidt, H. (2010). *Bankruptcy prediction: Static logit model versus discrete hazard models incorporating macroeconomic dependencies*. (Unpublished Master). Lund University, Lund.
- Fantazzini, D., & Figini, S. (2009). Random survival forests models for SME credit risk measurement. *Methodology and Computing in Applied Probability*, 11(1), 29-45.
- Feng, L., & Hendrikse, G. W. J. (2012). *Cognition and governance structure*. Ongepubliceerd manuscript.
- Fuertes, A., & Kalotychou, E. (2006). Early warning systems for sovereign debt crises: The role of heterogeneity. *Computational Statistics & Data Analysis*, 51, 1420-1441.
- Garicano, L., & Rayo, L. (2016). Why organizations fail: Models and cases. *Journal of Economic Literature*, 54(1), 137-192.

- Gordon, T. P., Fischer, M., Greenlee, J., & Keating, E. K. (2013). Warning signs: Nonprofit insolvency indicators. *International Research Journal of Applied Finance*, IV(3), 343-378.
- Greenlee, J. S., & Trussel, J. M. (2000). Predicting the financial vulnerability of charitable organizations. *Nonprofit Management & Leadership*, 11(2), 199-210.
- Hager, M. A. (2001). Financial vulnerability among arts organizations: A test of the Tuckman-Chang measures. *Nonprofit and Voluntary Sector Quarterly*, 30(2), 376-392.
- Hammond, T. H. (1994). Structure, strategy, and the agenda of the firm. In R. P. Rumelt, D. Schendel & D. J. Teece (Eds.), *Fundamental issues in strategy: A research agenda* (pp. 97-154). Boston: Harvard Business School Press.
- Heemskerk, K., Heemskerk, E., & Wats, M. (2014). Wat maakt goed onderwijsbestuur? De gedragsmatige determinanten van taakuitvoering door raden van toezicht in het voortgezet onderwijs. *Mens & Maatschappij*, 89(2), 151-174.
- Hendrikse, G. W. J. (1998). Screening, competition and the choice of the cooperative as an organisational form. *Journal of Agricultural Economics*, 49(2), 202-217.
- Hendrikse, G. W. J. (2004). Organisaties als systemen van attributen. *Tijdschrift Voor Economie En Management*, 49(3), 495-511.
- Honingh, Marlies, Kan, Jurriaan van, & Thiel, S. v. (2015). *Bestuurlijk handelen in relatie tot onderwijskwaliteit in het voortgezet onderwijs*. Ongepubliceerd manuscript.
- Inspectie van het Onderwijs. (2013). Financieel continuïteitstoezicht per 1 september 2013. Verkregen december/28, 2014, van <http://www.onderwijsinspectie.nl/actueel/nieuwsbrieven/details/financieel-continuïteitstoezicht-per-1-september-2013.html>
- Inspectie van het Onderwijs. (2009). *Sectorrapportage VO*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs. (2015). *Advies kengetallen*. Ongepubliceerd manuscript.
- Inspectie van het Onderwijs. (2016). Diverse inspectierapporten over de financiële positie van instellingen. Verkregen 04/29, 2016, van <http://www.onderwijsinspectie.nl/>
- Keating, E. K., Fischer, M., Gordon, T. P., & Greenlee, J. S. (2005). *Assessing financial vulnerability in the nonprofit sector*. (No. 27). Cambridge, Massachusetts: The Hauser Center for Nonprofit Organizations.
- Knippenberg, H., & Ham, W. v. d. (1994). *Een bron van aanhoudende zorg. 75 jaar ministerie van onderwijs [kunsten] en wetenschappen 1918-1993*. Assen: Van Gorcum.

- Lammers, J., Pelzer, B., Hendrickx, J., & Eisinga, R. (2007). *Categorische data analyse met SPSS : Inleiding in loglineaire analysetechnieken*. Assen: Koninklijke Van Gorcum.
- Lev, B., & Sunder, S. (1979). Methodological issues in the use of financial ratios. *Journal of Accounting and Economics*, 1(3), 187-210.
- Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, 8(4), 283-298.
- Ministerie van Onderwijs, Cultuur en Wetenschappen. (1963). *Wet op het voortgezet onderwijs (WVO)* [Wet van 14 februari 1963, tot regeling van het voortgezet onderwijs]
- Ministerie van Onderwijs, Cultuur en Wetenschappen. (1981). *Wet op het primair onderwijs (WPO)* [Wet van 2 juli 1981, houdende Wet op het basisonderwijs]
- Ministerie van Onderwijs, Cultuur en Wetenschappen. (2011). *Beleidsregel financieel toezicht po en vo 2011*
- Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109-131.
- Ritchie, W. J., & Kolodinsky, R. W. (2003). Nonprofit organization financial performance measurement. An evaluation of new and existing financial performance measures. *Nonprofit Management & Leadership*, 13(4), 367-381.
- Rodriguez, A., & Rodriguez, P. N. (2006). Understanding and predicting sovereign debt rescheduling: A comparison of the areas under receiver operating characteristic curves. *Journal of Forecasting*, 25(7), 459-479.
- Sah, R. K., & Stiglitz, J. E. (1986). The architecture of economic systems: Hierarchies and polyarchies. *The American Economic Review*, 76(4), 716-727.
- Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business*, 74(1), 101-124.
- Siau, C. (2009). De opstelling en het gebruik van een falingspredictiemodel. In D. Jaspers, & P. Vermoortel (Eds.), *Bedrijf & taal, opstellen voor wilfried Janssens/Business & language, essays in the honour of wilfried janssens* (pp. 3-26). Leuven: Peeters.
- Trussel, J. M. (2002). Revisiting the prediction of financial vulnerability. *Nonprofit Management & Leadership*, 13(1), 17-31.
- Tuckman, H. P., & Chang, C. F. (1991). A methodology for measuring the financial vulnerability of charitable nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, 20(4), 445-460.

4 Waardecreatie door verticale integratie tussen po, vo en mbo.²⁴

Samenvatting

Er zijn in Nederland rond 65 schoolbesturen (van de 1500) die verticaal zijn geïntegreerd: zij omvatten schoolsoorten in opvolgende levensfasen van leerlingen. Het is de vraag of dit een voordeel is omdat leerlingen soepel door het onderwijssysteem kunnen worden geleid en er efficiënter kan worden gewerkt, of een nadeel omdat de organisatie aanzienlijk complexer wordt. Gegevens rond leerlingenstromen, onderwijsresultaten en financiële resultaten wijzen uit dat verticale integratie van schoolbesturen meer nadelen dan voordelen biedt. Aan de wervingskant van leerlingen biedt het geen voordelen. Bijzondere concentraties in de instroom zijn eerder aan bijzondere denominaties en geografische omstandigheden toe te schrijven, dan aan een verticale structuur.

4.1 Inleiding

In het toezicht door de Inspectie van het Onderwijs (verder “de inspectie”) wordt multisectoraliteit als een risicofactor voor de kwaliteit en continuïteit beschouwd (Inspectie van het Onderwijs, 2015). Multisectoraliteit betekent dat schoolbesturen scholen hebben in meer dan een onderwijssector²⁵. De vaststelling van multisectoraliteit als risicofactor is vrij intuïtief tot stand gekomen. De overwegingen daarbij waren dat besturen met meer sectoren complexer in de aansturing zijn en zich minder kunnen concentreren op hun taken die zich over meer sectoren met verschillende wetgeving en verschillende omgeving moeten uitstrekken. De Commissie Onderzoek Financiële Problematiek Amarantis (2012) stelde dat omvang en complexiteit (onderwijs in meerdere sectoren) mede de oorzaak waren geweest van de financiële problematiek.

Multisectoraliteit is een vorm van verticale integratie, waarbij verschillende, opeenvolgende stappen in een productieproces worden samengebracht. Dat zou kunnen leiden tot bepaalde voordelen. In dat

²⁴ Dit hoofdstuk verschijnt in bewerkte vorm in nummer 1 van jaargang 28 van het Nederlands Tijdschrift voor Onderwijsrecht en Onderwijsbeleid (ntor).

²⁵ Het gaat om onderwijssectoren primair onderwijs, speciaal onderwijs, voortgezet onderwijs, middelbaar beroepsonderwijs en hoger onderwijs.

opzicht zou samenvoeging van verschillende sectoren niet als een risico hoeven te worden beschouwd. Als leerlingen, zoals hierboven gesteld, in een bepaalde beroepsstroom kunnen doorstromen, kan dat voordelig zijn voor de instelling (en de leerling). Anderzijds leidt multisectoraliteit tot een grotere complexiteit en het is de vraag of de voordelen van integratie opwegen tegen die problemen. Als dat niet het geval is, is multisectoraliteit terecht op te vatten als een risicofactor.

Om dat te onderzoeken wordt gekeken naar wat de economische theorie zegt over verticale integratie. Dit vindt plaats op basis van de transactiekostentheorie (Coase (1937), Williamson (1979)). Beschouwd wordt onder welke omstandigheden bedrijven kiezen voor vormen van verticale integratie en wat daarbij de voor of nadelen zijn. Dit hoofdstuk eindigt met een model waarmee die overwegingen systematisch kunnen worden getoetst. Vervolgens wordt bekeken hoe het in het Nederlandse onderwijs staat met verticale integratie binnen schoolbesturen. Daarbij komen aan de orde de bestaande stand van zaken en recente ontwikkelingen.

De integratie van schoolbesturen wordt vervolgens op twee manieren getoetst. Allereerst wordt gekeken of de vormen en overwegingen bij de integratie van opeenvolgende schoolsoorten aansluiten bij de theorie. Met andere woorden of de condities die volgens de theorie gelden voor integratie ook gelden bij de onderwijsbesturen. Vervolgens zal met kwantitatieve analyses worden onderzocht of de feitelijke integratie mogelijke voordelen oplevert. Het geheel wordt afgesloten met een conclusie en discussie.

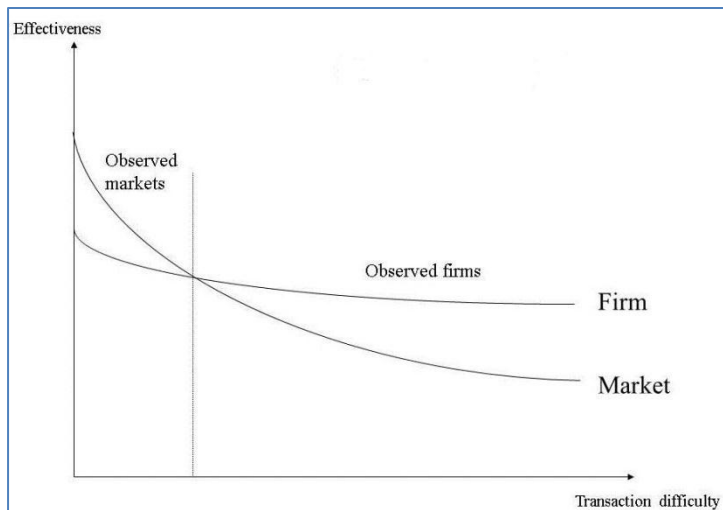
4.2 Verticale integratie

Pfeffer en Salancik (1978) beschrijven drie soorten operaties om de afhankelijkheid van andere organisaties te verkleinen, te weten verticale integratie, horizontale expansie en diversificatie of conglomeratie. Bij verticale integratie onderscheiden zij twee soorten, voorwaartse en achterwaartse integratie. Voorwaartse integratie betekent dat organisaties worden opgenomen die dicht bij de finale consument staan, om aldus de afhankelijkheid van bijvoorbeeld distributiekanaal te verkleinen. Achterwaartse integratie houdt in dat een organisatie uit een eerdere fase in de bedrijfskolom wordt opgenomen, bijvoorbeeld om de afhankelijkheid van grondstoffen te verkleinen. Horizontale expansie is erop gericht om organisaties met een vergelijkbaar aanbod op hetzelfde niveau in de bedrijfskolom op te nemen, om aldus de concurrentie te verkleinen. Diversificatie of conglomeratie tenslotte, is het verbinden met organisaties

die noch in dezelfde branche opereren, noch een directe leveranciers- of afnemersrelatie hebben.

Langlois en Robertson (1989) beschrijven vanuit de economische geschiedenis dat productie van origine in hoge mate geïntegreerd was en in de loop van de tijd steeds verder is opgedeeld. Waar origineel de productie van bijvoorbeeld drukwerk vanaf het maken van papier en inkt tot en met het verkopen van een boek in een hand was, is deze in de loop van de geschiedenis onderverdeeld in afzonderlijke stappen in het productieproces die door verschillende schakels werden verzorgd, waarbij de markt fungeerde als coördinerend mechanisme. Dit proces loopt echter niet onbeperkt in diezelfde richting door. De geschiedenis laat ook zien dat er met grote regelmaat organisatievormen voorkomen die de tegengestelde richting ingaan, en alleen al het persistente bestaan van bedrijven, die immers verschillende productiestappen integreren, lijkt in strijd met deze beweging.

Coase (1937) verklaarde dit verschijnsel uit de transactiekosten. Toenemende differentiatie leidt ertoe dat op steeds meer punten contracten tussen leverende en afnemende schakels in het proces moeten worden afgesloten. Zolang de aanname in de economische theorie was, dat deze contractering gratis plaatsvond, was er op theoretische gronden geen duidelijke reden voor het ontstaan en (blijven) bestaan van bedrijven. Coase liet evenwel zien dat als transactiekosten wel in de beschouwing worden betrokken, er eerder een punt ontstaat waarin het rendabel is de productie anders te organiseren dan via de markt. Daarbij bestaan transactiekosten niet alleen uit de kosten voor het zoeken van de juiste leveranciers of afnemers, het onderhandelen en het feitelijk afsluiten van contracten, maar ook uit fiscale randvoorwaarden. Die werken gunstiger bij leveringen binnen dan tussen bedrijven. Coase (1937) concludeerde dan ook dat als de kosten van interne coördinatie opwegen tegen kosten van externe transacties, het rendabel is activiteiten te bundelen in een bedrijf. Figuur 8 laat zien hoe de resultaten zich in respectievelijk bedrijven en op markten ontwikkelen naarmate de kosten van transacties toenemen. Rechts van de verticale lijn is het gunstiger de activiteiten in een bedrijf te organiseren dan in de vorm van externe transacties via de markt.



Figuur 8 Organisatiekosten versus (markt)transactiekosten (gebaseerd op Tadelis en Williamson, 2013)

Het zijn niet alleen direct aan transacties toe te wijzen kosten, zoals het verkennen van de markt, het voeren van onderhandelingen en het opstellen van contracten (bargaining costs, Hendrikse, 2003) die een belangrijke rol spelen bij de keuze voor verwerving (of levering) via de markt versus een vorm van integratie. Leveringen via de markt vergen een vorm van contractering. Daarbij is een onderscheid te maken tussen volledige en onvolledige contracten. Complete contractering veronderstelt dat alle relevante voorwaarden voor een contract kunnen worden vastgelegd en achteraf door een derde partij (de rechter) kunnen worden getoetst in geval van schending van de contractvoorwaarden. In veel situaties is complete contractering echter niet mogelijk omdat men niet alle voorwaarden kan overzien in verband met onzekerheden rond de transactiespecificaties. In dat geval kan het wenselijk zijn tot incomplete contractering over te gaan. Daarbij wordt slechts een deel van de afspraken formeel vastgelegd, en is er sprake van ruimte voor eigen interpretatie bij de contractpartijen van de afspraken en in een later stadium tot verschillende vormen van heronderhandeling (Hendrikse, 2003). Bij dit alles moet in aanmerking worden genomen dat de deelnemers aan de transacties beperkt rationeel zijn, risiconutraal en soms opportunistisch (Geyskens, Steenkamp en Kumar, 2006).

Williamson (1979) onderscheidt drie aspecten van transacties die bepalend zijn voor de keuze tussen marktorganisatie en bedrijfsorganisatie. De eerste daarvan

is de specificiteit van bedrijfsmiddelen (Mahoney, 1992). Hiermee wordt bedoeld dat een of beide partijen bijzondere investeringen moet doen die in overwegende mate bruikbaar zijn in het kader van de onderlinge relatie, maar die buiten die relatie veel minder waard zijn. Het verschil tussen beide wordt ook wel aangeduid met "sunk costs" (Hendrikse, 2003). Een voorbeeld is het aanleggen van speciale haveninstallaties om leveranties tussen de twee partijen te faciliteren, die niet zonder meer bruikbaar zijn in andere relaties, maar ook specifieke investeringen in scholing van personeel kunnen leiden tot sunk costs. Dergelijke investeringen kunnen vooral bij een verlenging van het contract tot onzekerheid leiden. De investerende partij kan dan in een nadelige positie komen te verkeren, omdat de andere partij gunstiger voorwaarden kan bedingen, wetend dat de investerende partij een verlies van de sunk costs zal willen vermijden als de relatie beëindigd zou worden (Hendrikse, 2003).

Een tweede categorie van onzekerheden ligt rond transactiespecificaties. Deze onzekerheden kunnen van verschillende aard zijn. Geyskens, Steenkamp en Kumar (2006) onderscheiden de volgende vormen van onzekerheid. Van omgevingsonzekerheid is sprake als de omstandigheden rond de transactie te onvoorspelbaar zijn om vooraf in een contract te kunnen worden afgedekt. In dergelijke gevallen moeten tussentijds veel kosten gemaakt worden om tot aanpassingen van het contract te komen. Gedragsonzekerheid betekent dat de prestaties achteraf niet gemakkelijk zijn vast te stellen, zodat onduidelijk is of aan de afspraken is voldaan. Verder kan er sprake zijn van volumeonzekerheid waarbij vooraf de af te nemen hoeveelheden onduidelijk zijn, hetgeen kan leiden tot overschotten of juist tekorten aan capaciteit. Tenslotte is er technologische onzekerheid, waarbij de specificaties van te leveren componenten aan verandering onderhevig kunnen zijn.

Barzel (1982) benoemt meetkosten als een aparte categorie die in ogenschouw zou moeten worden genomen. Meetkosten zijn alle kosten die te maken hebben met onzekerheid rond de vraag of geleverd wordt waarvoor wordt betaald. Dat betreft niet enkel technische kosten om metingen uit te voeren, maar ook de kosten van (herstel van) fouten en van het zoek- en onderhandelingsproces rond de transactie. De meet- en herstellkosten sluiten nauw aan bij de hierboven genoemde gedrags- en volumeonzekerheden. De laatste overweging die Barzel noemt, vertoont een nauwe verwantschap met de bargaining costs (Hendrikse, 2003).

Een derde categorie is de frequentie van de transacties (Geyskens e.a. 2006). Williamson (1979) acht het plausibel dat frequente transacties goedkoper binnen een organisatie kunnen plaatsvinden dan in de markt

en dat dat dus een argument voor integratie zou kunnen zijn. Geyskens e.a (2006) stellen echter vast dat er weinig empirisch onderzoek naar is gedaan en dat dat geen duidelijke correlaties laat zien. Tegenover de vermindering van de transactiekosten in de markt, kunnen binnen een organisatie andere kosten, zoals influence costs (Hendrikse, 2003) komen staan. Daarop zal verderop nader worden ingegaan. Frequentie transacties vormen wel een geschikt uitgangspunt voor relationele contracten (Baker, Gibbons en Murphy, 2002) die zowel intern binnen een bedrijf als tussen bedrijven kunnen worden afgesloten. Dergelijke contracten functioneren in langdurige relaties tussen betrokkenen, waarbij sprake is van wederzijds vertrouwen. Baker e.a. (2002) vinden dat daarbij managers een belangrijke rol spelen in het onderhouden van de relatie.

Relationele contracten kunnen minstens voor een deel voorzien in de onzekerheden die samenhangen met incomplete contractering. Daarbij is het onmogelijk precies de condities rond het contract vast te leggen of, in het geval van specifieke investeringen, is het onzeker of een contractpartij in de toekomst, bij verlenging van het contract, opportunistisch gedrag aan de dag gaat leggen. De kwestie die hierbij een rol speelt, is die van de toekenning "residual rights". Grossman en Hart (1986) onderscheiden bij de contractering twee rechten: "specific rights" en "residual rights". Specific rights zijn al die afspraken die precies in een contract worden geregeld, bijvoorbeeld investeringen in een gezamenlijke relatie en de verdeling van opbrengsten uit die relatie. Residual rights zijn al die zaken waarvoor dergelijke precieze afspraken niet mogelijk zijn. Zij kunnen betrekking hebben op onzekere winsten of verliezen. Grossman en Hart laten zien dat de verdeling in dergelijke gevallen beïnvloed kan worden door de governance van de samenwerkingsrelatie en de mogelijkheden tot opportunistisch gedrag.

De te kiezen vorm van governance moet, gegeven de onzekerheden, leiden tot een totaal dat maximale complementariteit vertoont (Hendrikse, 2004). Er zijn verschillende andere vormen van governance mogelijk die als kenmerken hebben dat er sprake is van verschillende combinaties van de elementaire rechten in een governancestructuur: eigendomsrechten, inkomensrechten en beslissingsrechten (Hendrikse, 2003, Hendrikse en Feng, 2012). Baker, Gibbons en Murphy (2008) hebben een aantal van de mogelijke allianties theoretisch onderzocht op basis van een analyse van netwerken tussen een groot aantal farmaceutische en biotechnische bedrijven tussen 1973 en 2001. Het model geeft voor een aantal verschillende configuraties van toewijzing van eigendomsrechten, beslissingsrechten en inkomensrechten de te verwachten resultaten voor de betrokken partijen.

De meest vergaande vorm van integrerende governance is de overname van de contractpartij. Dit is wat feitelijk onder verticale integratie wordt verstaan. De volledige inkomens- en beslissingsrechten gaan daarbij over op de overnemende partij. De eerst zelfstandige onderneming wordt een onderdeel van de andere. Tegenover het voordeel van onzekerheidsreductie kan een dergelijke configuratie ook nadelen hebben. Zo kan de gerealiseerde situatie minder efficiënt zijn omdat de tucht van de markt voor de overgenomen partij verdwijnt (Baker e.a. 2008). Bovendien wordt daarbij de ene soort kosten (bargaining costs, haggling costs) vervangen door een andere soort (influence costs, costs of bureaucracy) (Powell, 2015). Powell laat zien dat deze soorten kosten dezelfde oorsprong hebben en geminimaliseerd kunnen worden door de juiste vorm van governance. Dat kunnen tussen de extremen van governance van de markt en volledige verticale integratie verschillende alternatieve vormen zijn. Hendrikse (2003) illustreert aan de hand van de specificiteit van de investeringen, dat als deze hoog zijn, de governancekosten van een hiërarchische structuur lager uitkomen dan bij governance via de markt. Deze laatste komen gunstiger uit de bus bij een laag niveau van specificiteit.

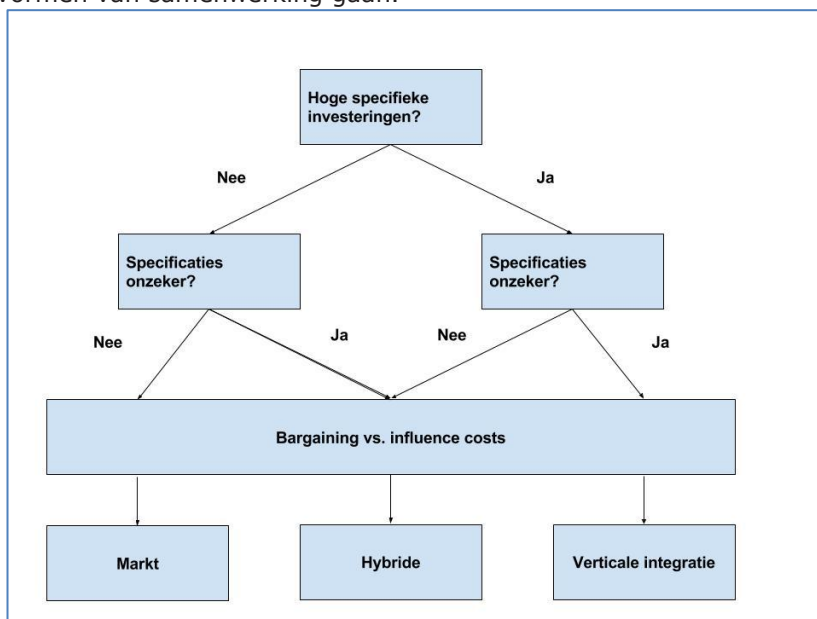
In het gebied tussen de markt en volledige integratie kunnen zeer verschillende vormen voorkomen van wat algemeen wordt aangeduid als hybride vormen (Ménard, 2012). Ménard onderscheidt, op basis van de vorm van governance, in dat tussengebied drie categorieën van hybriditeit. De eerste categorie, die het dichtst aanligt tegen de markt, bestaat uit "information-based-networks" waarbij betrokken organisaties in hoge mate zelfstandig blijven maar intensief kennis uitwisselen, vaak via gezamenlijke platforms. Eigenaarschap over de bedrijfsmiddelen en decision rights blijven bij de organisaties. Dergelijke verbindingen zijn onder meer te vinden in de biotechnologie en computertechnologie (Baker e.a., 2008).

Het dichtst tegen verticale integratie aan liggen hybride vormen die Ménard (2012) typeert als "strategic center". Hierbij is sprake van een strakke coördinatie van de activiteiten van betrokken partners. Hieronder vallen bijvoorbeeld joint-ventures waarbij belangrijke delen van het eigenaarschap en de decision rights overgebracht zijn. Daartussenin is sprake van een gebied met "third party coordination". Daarbij houden de partners controle over hun eigen bedrijfsmiddelen maar is er sprake van gezamenlijke investeringen. De controle daarop wordt in handen gelegd van een derde partij.

Om te bepalen welke vorm van governance geschikt is in een bepaalde context, zijn deze overwegingen hieronder in een classificatieschema

gezet in Figuur 9. De primaire keuze is erin gelegen of er al dan niet sprake is van hoge specifieke investeringen. Deze keuze komt uit de literatuur als vrij dominant naar voren en deze lijkt relatief eenvoudiger vast te stellen dan de andere. Naar de mate waarin er meer sprake is van specifieke investeringen, zal de keuze naar een meer geïntegreerde vorm gaan. De secundaire keuze is die van de onzekerheid rond de transactiespecificaties. Is die onzekerheid groter, dan ligt het meer voor de hand richting een geïntegreerd vorm te gaan.

De derde keuze is het minst eenduidig. Daarin vindt een afweging plaats tussen enerzijds de kosten die gemaakt moeten worden in marktomstandigheden om tot transacties te komen (bargaining of haggling costs) tegenover anderzijds de kosten die in een organisatie gemaakt moeten worden (influence costs, bureaucracy costs) (Powell, 2015). Naar de mate waarin de ene soort kosten de andere overheerst, zal de keuze meer richting de markt of juist richting meer geïntegreerde vormen van samenwerking gaan.



Figuur 9 Classificatie Governancevormen

Dit classificatieschema zal in de analyse verder als referentie worden gebruikt bij het beschouwen van de integratie van onderwijsinstellingen.

Over verticale integratie in institutioneel opzicht in het onderwijs, bestaat weinig literatuur. De literatuur die er internationaal wel is, is vooral gericht op samenwerkingsverbanden tussen organisaties. Wat men hierbij

verticale integratie noemt, wijkt daarmee af van de definitie die hierboven in de economische theorie is gehanteerd. Er is namelijk geen sprake van verticaal eigenaarschap, maar veeleer van een inhoudelijke afstemming van activiteiten in samenwerkingsverbanden in de vorm van een relationeel contract. In onderwijskundige termen worden in die samenwerkingsverbanden doorlopende leerlijnen gecreëerd. Die doorlopende leerlijnen worden verzorgd door verschillende instanties die daarvoor samenwerkingsovereenkomsten hebben afgesloten. Er is hier geen sprake van hoge, specifieke investeringen en inspanningen en output zijn slecht te meten tussen de participanten. Dit leidt dan tot relationele contracten tussen organisaties (Baker e.a., 2002), een hybride vorm van governance, als de meest voor de hand liggende organisatievorm en daar is in de praktijk ook sprake van. In termen van Ménard (2012) zouden zij als information-based networks te omschrijven zijn.

Zo beschrijven Rosenthal, Worley, Mugford en Stagg (2004) en Dick, King, Mitchell, Kelly, Buckley en Garside (2007) en Thomson en Anderson (2009) de ontwikkeling van verticaal geïntegreerde medische opleidingen in zeer dunbevolkte, landelijk gebieden in Australië. Het is in die gebieden onmogelijk opleidingen in niveaus te scheiden door hun beperkte omvang. Als studenten naar universiteiten in stedelijke gebieden gaan, komen ze meestal niet meer terug naar het platteland. Bovendien vereist geneeskunde op het platteland een brede opleiding en ervaring in plaats van een smallere, specialistische benadering die in dichtbevolkte gebieden mogelijk is. De studenten doorlopen een gekoppeld traject dat zich deels op stedelijke universiteiten en deels in de lokale praktijk (General Practice) afspeelt. Er is daarbij echter geen sprake van institutionele integratie, maar van samenwerkingsverbanden tussen de verschillende deelnemers in de vorm van relationele contracten.

4.3 Verticale integratie in het onderwijs in Nederland

In het bekostigd onderwijs in Nederland is sprake van een zeer divers geheel aan bestuurlijke verbanden. Kenmerkend voor het Nederlandse onderwijs is de onderverdeling naar denominaties. Het belangrijkste onderscheid is dat tussen het openbaar en het bijzonder onderwijs. Het eerste wordt aangestuurd door de overheid (vanouds rijk en gemeenten), het laatste door private rechtspersonen, vroeger vaker verenigingen maar tegenwoordig in toenemende mate stichtingen. Het bijzonder onderwijs kent op zijn beurt weer een groot aantal denominaties, variërend van scholen op verschillende confessionele grondslagen tot scholen op een

algemeen bijzondere (niet-confessionele) basis. Hierbij zijn vaak weer scholen te vinden met een bepaalde pedagogische-didactische visie (DUO, 2015a). De verdeling naar denominaties is al gedurende langere tijd nagenoeg onveranderd. In het basisonderwijs is een derde van de scholen openbaar. Protestants-christelijk onderwijs in verschillende stromingen en katholiek onderwijs omvatten elk 30 procent. Het resterende deel bestaat uit overige, bijzondere scholen van oftewel algemene signatuur of kleine, confessionele stromingen (OCW, 2014).

Het aantal scholen en schoolbesturen is al enkele decennia aan het dalen (Onderwijsraad, 2008). Dit is het gevolg van het beleid van het ministerie van OCW in die periode, waarin steeds meer werd gedecentraliseerd. Schoolbesturen hebben in Nederland een zeer grote mate van autonomie en die is sinds de jaren negentig nog fors toegenomen. Binnen algemene richtlijnen voor de inrichting van het onderwijs, bevoegdheden van personeel en de besteding van middelen zijn zij zelf verantwoordelijk voor hun onderwijskundige aanpak, bedrijfsvoering, personeelsbeleid, investeringen en dergelijke. Daartoe ontvangen zij bekostiging van het ministerie.

Om in staat te zijn de gevolgen van de grotere verantwoordelijkheid het hoofd te bieden, waren scholen, mede gestimuleerd door het ministerie, op grote schaal gefuseerd om voldoende bedrijfsmatig volume te krijgen. Dat betekende overigens lang niet altijd dat scholen zelf aanzienlijk groeiden, maar dat zij bestuurlijk fuseerden en onder een bredere paraplu kwamen. Hierbij kunnen ook denominaties in het gedrang komen. Dat is bijvoorbeeld te zien aan de ontwikkeling van het aantal samenwerkingsbesturen, waarin scholen van meerdere denominaties onder een bestuur zijn samengebracht. Omdat de politiek bang was voor uitwassen, is per 1 oktober 2011 de Wet fusietoets (OCW, 2011) van kracht geworden. Daarmee moeten besturen hun voorgenomen fusies voorleggen aan de minister die daarover beslist na advies door de Commissie Fusietoets.

Niettemin is het aantal scholen en besturen in het primair onderwijs blijven dalen, waarbij de terugloop van het aantal leerlingen eveneens een belangrijke rol speelt. Tussen 2009 en 2013 daalde het aantal basisscholen met ruim drie procent van 6881 naar 6649 (OCW, 2014). Het aantal besturen is in die periode eveneens gedaald met acht procent en de gemiddelde bestuursgrootte, uitgedrukt in aantal scholen, is dienovereenkomstig gestegen. In het voortgezet onderwijs is het aantal scholen over de hele periode stabiel terwijl het aantal leerlingen er nog licht steeg. In het mbo is het aantal instellingen licht toegenomen, maar dat is veroorzaakt door het opknippen van Amarantis.

Deze ontwikkeling speelt zich vooral in de horizontale sfeer af (Commissie Fusietoets Onderwijs (CFTO), 2015), dus tussen scholen van dezelfde onderwijssoort. Fusies in de verticale sfeer zijn de afgelopen jaren beperkt in aantal. Overwegingen van verticale verbindingen kunnen in het onderwijs gelegen zijn in het scheppen van doorlopende leerwegen, speciaal in het beroepsonderwijs. De Onderwijsraad (2009) pleit in zijn verkenning over de ontwikkelingen in het middelbaar beroepsonderwijs voor wat genoemd wordt verticalisering. Ononderbroken leerwegen van vmbo naar mbo in verwante branches (bijvoorbeeld techniek of zorg) zouden kunnen bijdragen aan het verminderen van de schooluitval van met name de wat zwakkere leerlingen. Dergelijke trajecten (VM2-trajecten genaamd) worden in de regel vormgegeven in samenwerkingsverbanden tussen vmbo-scholen en ROC's en leiden niet tot een bestuurlijk samengaan van de instellingen (Amarantis was daarop een uitzondering). In die zin is er dus geen sprake van verticale integratie maar van een nauwere samenwerking tussen leverancier (vmbo) en afnemer (ROC).

De Onderwijsraad (2008) zegt in vrij algemene termen over verticalisering en diversificatie:

"Daarnaast kunnen zich in theorie verticale schaalvoordelen voordoen bij fusies tussen verschillende onderwijssectoren (bijvoorbeeld primair en voortgezet onderwijs) door lagere transactiekosten. Daar staat tegenover dat verticale integratie kan leiden tot vermindering van de flexibiliteit en een zwakkere prikkel tot innovatie.

Tot slot kunnen scopevoordelen worden behaald door diversificatie, bijvoorbeeld door fusies met andere maatschappelijke organisaties (zoals kinderopvang of welzijn). Op die manier kan een organisatie risico's afdekken (bijvoorbeeld van fluctuerende leerlingenaantallen) of op een goedkope manier combinaties van dienstverlening aanbieden (bijvoorbeeld naschoolse opvang).

In de praktijk blijken dergelijke sectoroverschrijdende organisaties vaak moeilijk bestuurbaar te zijn, waardoor extra coördinatiestructuren nodig zijn. In het algemeen lijken de kosten van diversificatie niet op te wegen tegen de opbrengsten."

Waterreus (2009) merkt over verticale fusies op dat die in theorie schaalvoordelen kunnen opleveren maar dat verticale integratie, naast de ook al door de Onderwijsraad (2008) genoemde, verminderde prikkel tot innovatie, ook kan leiden tot vermindering van de flexibiliteit en keuzevrijheid. Het zou voor ouders en deelnemers gemakkelijker worden door te stromen tussen scholen die vallen onder hetzelfde schoolbestuur en de drempel om over te stappen zou groter worden. Gelijktijdig stelt hij

vast dat het aantal verticale besturen in het basis- en voortgezet onderwijs aanzienlijk is afgenomen.

Dit laatste is waarschijnlijk het gevolg van de bestuurlijke verzelfstandiging van het openbaar onderwijs, die vooral rond 2007 een grote vlucht nam (CBS, 2008). Waar dat openbaar onderwijs eerst rechtsreeks onder het gezag van de gemeente viel, en daarmee automatisch verticaal was als er openbaar primair en voortgezet onderwijs in die gemeente was, zijn veel openbare scholen onder het bestuur van zelfstandige rechtspersonen gebracht. In sommige gevallen is al het onderwijs bij een (verticaal) bestuur ondergebracht, terwijl er in andere gemeenten voor gekozen is dit te doen in afzonderlijke besturen voor primair en voortgezet onderwijs. Tabel 4.1 laat ook zien dat de teruggang uitsluitend zit in de combinatie po/vo (Bao/vo genaamd in de tabel), waar gemeenten vaak het bevoegd gezag over waren. Waterreus concludeert uit de ontwikkeling van het aantal verticale besturen dat er bij politici en onderwijsbestuurders kennelijk twijfels bestaan over de meerwaarde van dergelijke verbanden. Na 2008 is het aantal verticale besturen in primair en voortgezet onderwijs, blijkens de gegevens van DUO (2015) wel weer gestegen.

Tabel 4.1: Ontwikkeling van het aantal schoolbesturen in het onderwijs 2000-2008 (uit: Waterreus, 2009)

	2000	2004	2008
Bao (incl. sbao en (v)so)	1877	1471	1183
Bao/vo	74	31	21
Vo	316	287	278
Vo/mbo	28	27	27
Mbo	42	37	37
Mbo/hbo	1	1	3
Hbo	52	41	35
Wo	13	13	13

Tabel is gebaseerd op OCW (2009). Kerncijfers, 2004-2008. Den Haag: Ministerie van OCW. Inclusief de instellingen voor groen onderwijs (vallend onder categorie vo/mbo). Besturen voor speciaal basisonderwijs en (voortgezet) speciaal onderwijs zijn niet apart vermeld maar meegeteld bij de besturen van het basisonderwijs (bao).

In het beroepsonderwijs komen relatief de meeste vormen van verticale integratie voor. De geschiedenis van het beroepsonderwijs is vanouds nauw verbonden met het bedrijfsleven en het is pas veel later dan het funderend onderwijs door de overheid gereguleerd, waarbij er steeds ook sprake is geweest van onderwijs vanuit de bedrijven zelf. Moerkamp en Onstenk (1999) beschrijven hoe het zich in de negentiende eeuw ontwikkelde vanuit particulier initiatief en vanaf 1921 werd geregeld in de Nijverheidsonderwijswet. De Mammoetwet bracht vanaf 1968 een

duidelijke samenhang en structuur aan tussen de uiteenlopende vormen en niveaus van het beroepsonderwijs en die van het algemene onderwijs. In de jaren tachtig en negentig heeft dat verder gestalte gekregen in afzonderlijke wetgeving van het hbo en mbo.

De algemene tendens van de ontwikkeling is, net als Langlois en Robertson (1989) beschrijven voor de economie als geheel, er een van toenemende differentiatie en specialisatie. Daarin hebben zich in de loop van de laatste eeuw verschillende bewegingen voorgedaan, soms in de richting van (re)integratie, soms naar verdere opsplitsing. De overwegingen waren daarbij niet altijd op de inhoud of de bedrijfsvoering gericht maar vielen meer samen met bredere stelselwijzigingen in het onderwijs. Een drietal voorbeelden laat dat treffend zien.

Een illustratief voorbeeld van de ontwikkeling in de sector die momenteel het sterkst geverticaliseerd is, is die van de Agrarische Opleidingscentra (AOC's). Kooij (2008) beschrijft het traject van ruim twintig jaar waarin uiteindelijk rond 1994 de verticale AOC's waren gevormd. Daarbij speelden niet in de eerste plaats pedagogische of bedrijfseconomische opvattingen over de voordelen van een verticale structuur een rol, maar primair politieke. Een verticale structuur zou belangrijke argumenten leveren om het agrarisch onderwijs binnen de eigen kring (het ministerie van Landbouw en Visserij) te houden waar het onderwijs begin jaren zeventig nogal verspreid was en vaak gekoppeld zat aan overige vormen van beroepsonderwijs die weer onder het ministerie van Onderwijs vielen. Ook opvattingen over de mogelijke ontwikkeling van de middenschool in de jaren zeventig, die zouden leiden tot veralgemening van het groene onderwijs, waren daarbij van belang. Aangezien er geen sprake is van hoge, specifieke investeringen, zou hierbij een relationeel contract meer voor de hand hebben gelegen²⁶ dan een institutioneel samengaan.

Juist de tegengestelde beweging vond plaats in het laboratoriumonderwijs (Van der Willik, 2010). Dat was van oorsprong verticaal georganiseerd in mbo-hbo-opleidingen in laboratoriumscholen. Door twee grote operaties, eerst in het hbo en later in het mbo zijn deze opleidingen van elkaar gescheiden en opgegaan in horizontale verbanden. Dit heeft geleid tot een reductie van opleidingen en uitstroom van studenten, speciaal in het mbo. Van der Willik pleit dan ook voor een herschikking in hiërarchische verbanden. Het is echter de vraag of dat, in termen van de classificatie in Figuur 9 voor de hand ligt. Er is weliswaar sprake van hoge investeringen

²⁶ Voor een deel van die investeringen, met een zeer specifiek karakter, kende het Agrarisch Onderwijs daarnaast een aparte verzorgingsstructuur, de Innovatie- en praktijkcentra, waar scholen dergelijke voorzieningen konden inhuren. Daarvoor was institutioneel samengaan dus niet nodig.

in apparatuur die zowel in het mbo als het hbo bruikbaar is (en waarschijnlijk in sterker mate dan in het groene onderwijs), maar dit zijn geen investeringen die specifiek op het faciliteren van de verticale koppeling gericht zijn. Ook hier zouden relationele contracten in de behoeften moeten kunnen voorzien. Dit is ook iets wat Van der Willik als alternatieve optie benoemt, aangezien een echte doorbreking van de structuur ingewikkelde wetswijzigingen zou vereisen.

Een derde ontwikkeling heeft zich voorgedaan in het onderwijs in de gezondheidszorg. Tot 1996 was dit voor een belangrijk deel ingericht als inservice-onderwijs. Instellingen voor gezondheidszorg namen personeelsleden aan die binnen het kader van hun werk werden opgeleid. Daarnaast waren er algemenere opleidingen op mbo- en hbo-niveau die leerlingen opleidden. De inservice-opleidingen werden veelal verzorgd door of eigen opleidingsinstituten van de zorginstellingen (volledige integratie) of door een opleidingsschool voor een groep zorginstellingen (lange-termijncontract, joint venture). Het bezwaar bij veel van die opleidingen was dat ze vaak zeer specifiek op de eigen instelling waren gericht waardoor leerlingen te weinig mogelijkheden hadden om naar andere instellingen over te stappen. In 1996 is daarom besloten het onderwijs in de gezondheidszorg volledig van de instellingen los te halen en te plaatsen bij het ministerie van OCW (Commissie Kwalificatiestructuur Verpleging en Verzorging, 1996; Abu-Saad Huijer, 1996).

Fusies in het onderwijs zijn sinds 2011 onderworpen aan de Wet Fusietoets die als gevolg heeft dat de fusie moet worden voorgelegd aan de Commissie Fusietoets Onderwijs (CFTO). De openbare rapporten van de CFTO bieden inzicht in de overwegingen die schoolbesturen hebben bij fusies. Het overgrote deel van die fusies is horizontaal, dus tussen schoolbesturen met scholen in dezelfde onderwijssector. Het aantal intersectorale fusies is betrekkelijk gering (16 van 2012 tot en met 2014) en ze zijn ook niet altijd verticaal. Meer dan de helft van de fusies betreft die tussen schoolbesturen voor speciaal onderwijs (vallend onder de Wet op de Expertisecentra) en primair onderwijs (vallend onder de Wet op het Primair Onderwijs). Het gaat daarbij niet om leerlingen in opeenvolgende levensfasen, maar om leerlingen die bijzondere onderwijsvoorzieningen nodig hebben en dus is er eigenlijk geen sprake van verticale integratie.

Slechts in zes gevallen was echt sprake van een verticaal bestuur, waarbij dat in twee gevallen ook voordien al het geval was (het ontvangende bestuur had al meer sectoren). De overwegingen die daarbij golden waren deels van ideologische aard: het waarborgen van een doorlopende lijn volgens de eigen denominatieve of didactische overtuiging. Voor een

ander deel waren zij onderwijskundig onderbouwd, namelijk met het aanbod van doorlopende leerlijnen en een doorlopend begeleidingstraject. Enkele keren werden ook overwegingen van schaalvoordelen en efficiency genoemd en in een geval pure overleving van een bestuur. Dat verschillende overwegingen nogal door elkaar lopen, blijkt ook uit de beschrijving van een verticale fusie van voor de CFTO (Naaijens, 2010). Daarin wordt in de onderbouwing gestart met het verkleinen van de kloof tussen po en vo, worden allerlei onderwijskundige en schaalvoordelen benoemd, maar blijkt de belangrijkste overweging toch de denominatie te zijn.

4.4 Onderwijsintegratie in de economische theorie

Een essentieel verschil tussen de omgeving waarin de bedrijven in de beschreven economische theorie opereren en die waarin het onderwijs in het algemeen opereert, is dat er geen sprake is van (willoze) producten die verschillende stadia van hun ontwikkelingsproces doorlopen, maar van leerlingen die in de loop van hun ontwikkeling verschillende vormen van dienstverlening krijgen waarmee zij een transformatie ondergaan. Daarnaast is er sprake van een belangrijk verschil ten aanzien van het formele eigendom van bekostigde (niet op winst gerichte) onderwijsinstellingen versus ondernemingen.

De keuzevrijheid van leerlingen en hun ouders is in het Nederlandse onderwijs ook wettelijk geregeld (Onderwijsraad, 2008). Deze keuzevrijheid brengt met zich mee dat relaties tussen scholen in opeenvolgende sectoren nooit gebaseerd kunnen zijn op leverings- of afnameverplichtingen van leerlingen aan of van afnemende of toeleverende instellingen. Dat betekent dat er voor de partners in de samenwerkingsrelatie geen zekerheid kan worden verschaft over input en output in hun relatie. Bovendien is er sprake van bijzondere marktomstandigheden. Het onderwijs wordt volledig bekostigd door de overheid. Dat gebeurt onder een aantal voorschriften die beperkingen opleggen aan de mogelijkheden van bedrijfsmatig opereren.

Bekostigd onderwijs is niet gericht op het maken van winst. Als residual rights dus betrekking zouden hebben op de (moeilijke) verdeling van de winst die door integratie vereenvoudigd zou worden, zal daar in het Nederlandse onderwijs maar beperkt sprake van kunnen zijn. Dat zou het geval kunnen zijn als efficiencywinsten kunnen worden geboekt die dan weer op een andere manier aan het onderwijs ten goede komen. Door

gebruik te maken van gemeenschappelijke infrastructuur of personeel kunnen twee instellingen goedkoper opereren en de verdiensten die dat oplevert kunnen niet gemakkelijk aan beide partijen worden toegerekend. Integratie kan hier een oplossing voor de residual rights bieden. Of dergelijke efficiencyvoordelen zich daadwerkelijk voordoen, zal verderop in de kwantitatieve analyse aan de orde komen.

Een andere vorm van winst zou erin gelegen kunnen zijn, ook al is er geen sprake van leveringsverplichtingen, dat de doorstroom van leerlingen binnen het eigen bestuur door de verticale verbanden wel wordt bevorderd. Dat zou dan leiden tot omzetvergroting binnen het eigen bestuur aangezien de bekostiging van het funderend onderwijs volledig is gebaseerd op het leerlingenaantal. Het kan ook omzetverhogend werken als deze leerlingen, beter geëquipeerd, doorstromen van vmbo naar mbo en daarbij betere eindresultaten boeken. De bekostiging van het mbo is namelijk naast op leerlingenaantal, deels gebaseerd op diplomarendement. In de kwantitatieve analyse zal ook hiernaar verder onderzoek plaatsvinden.

Een belangrijke overweging voor verticale integratie is de noodzaak specifieke investeringen in activa of personeel te doen om aldus de relatie tussen de opeenvolgende schakels vorm te geven. Van dergelijke investeringen is in het Nederlandse onderwijs alleen al daarom geen sprake, dat leerlingen vanuit zeer uiteenlopende scholen vervolgonderwijs kunnen doen. Deze scholen hoeven daartoe geen andere investeringen te doen dan die welke nodig zijn om het reguliere eindniveau te bereiken. Docenten hebben voor de verschillende sectoren verschillende opleidingen²⁷ en ook de materiële voorzieningen zijn verschillend per sector. Dat geldt al helemaal voor het po en vo, maar ook in situaties als het groene onderwijs of het laboratoriumonderwijs komen dergelijke investeringen niet voor. Het is best denkbaar, zoals Willik (2010) ook stelt, dat er schaalvoordelen kunnen worden gerealiseerd in het gezamenlijk gebruik van investeringen in opeenvolgende schakels, maar die investeringen zijn zelf niet specifiek voor een eventuele, geïntegreerde relatie. Ook hier geldt dat leerlingen niet noodzakelijk door hoeven stromen in de beroepskolom, maar ook vanuit andere opleidingen kunnen instromen.

²⁷ In een van de fusieaanvragen bij de CFTO (2015) is sprake van een verbinding op een heel specifieke, onderwijskundige richting (Vrije School). Daarbij zou specifieke scholing van docenten een rol kunnen spelen in het verticale verband. Voor het overige moeten docenten in po en vo verschillende soorten onderwijsbevoegdheden hebben.

Kijkend naar de classificatie (Figuur 9) laten deze overwegingen zien dat er op theoretische gronden weinig of geen reden is om scholen verticaal te integreren. Allereerst is er geen sprake van specifieke investeringen, zodat institutionele integratie eigenlijk al niet meer noodzakelijk is. Dat betekent dat enkel marktvormen of hybride vormen voor de hand liggen. Sturing van resultaten kan slechts in beperkte mate plaatsvinden, onder meer door de keuzevrijheid van ouders en leerlingen voor een vervolgschool, die niet gebonden is door een verticale, institutionele verbinding tussen scholen en het bestuur.

Verder zijn de specificaties niet heel onzeker, met inachtneming van de individuele verschillen die er bij leerlingen natuurlijk altijd zijn. Aan het eind van het primair onderwijs is er sprake van een schooladvies, veelal gecombineerd met een gestandaardiseerde toets waaraan de vaardigheden van leerlingen op enkele kernterreinen zijn af te lezen. Deze toets kan mogelijk negatieve externaliteiten, het geven van te gunstige adviezen waarmee problemen op het vervolgonderwijs worden afgewenteld, inzichtelijk maken. Het voortgezet onderwijs kent een landelijk examen. Voor vervolgonderwijs vormt dat een belangrijke mogelijkheid om leerlingen op hun capaciteiten te kunnen taxeren. Alleen als de individuele verschillen van leerlingen buiten deze kernonderdelen van grote relevantie zijn voor de vervolgopleiding, zou een vorm van verticale verbinding een voordeel kunnen bieden.

Een laatste overweging is die tussen enerzijds de kosten in de markt (bargaining) versus de kosten in de organisatie (influence, bureaucracy). Het gegeven dat een belangrijk deel van de leerlingen altijd naar andere scholen gaat of van andere scholen komt dan die van het eigen bestuur, brengt met zich mee dat in de wervingssfeer maar betrekkelijk weinig winst kan worden gemaakt. Deze kosten blijven voor een deel van de leerlingen hoe dan ook bestaan. Gelijktijdig komt er in een verticaal geïntegreerde organisatie extra complexiteit bij, deels door het groter worden van de organisatie maar vooral omdat er een compleet stelsel van context en wetgeving wordt toegevoegd.

Op theoretische gronden lijkt er daarom weinig reden te zijn om tot verticale verbanden over te gaan. Aangezien er door uiteenlopende oorzaken toch sprake is van verticaal geïntegreerde instellingen, kan het nuttig zijn te onderzoeken of daar niettemin voordelen mee kunnen worden gehaald. Dat zal gebeuren in de navolgende, kwantitatieve analyse waarin zal worden beschouwd of dat het geval is op het vlak van wervingsvoordelen, onderwijsresultaten, financiële prestaties en overheadkosten.

4.5 Vraagstelling

Het geheel aan vormen van en motieven achter verticale integratie of juist differentiatie in het Nederlandse onderwijs laat een bonte verscheidenheid zien. De vraag bij dit alles is dan ook of de vormen van verticale integratie nu voordelen opleveren, dan wel terecht door de inspectie als een risico worden geduid (en ook beide is natuurlijk mogelijk). De voordelen van verticale integratie hoeven niet primair op het financiële vlak te liggen. Naast de moeilijker meetbare ideologische voordelen kunnen ook leerlingen er baat bij hebben volgens een dergelijke lijn te worden opgeleid, en dat zou tot uitdrukking kunnen komen in hun studieresultaten. De positie aan de wervingskant kan worden versterkt doordat er een grotere kans is op instroom vanuit de eigen organisatie.

Aan de financiële of de bestuurlijke kant kan verticale integratie leiden tot risico's. De schaal van de totale instelling neemt weliswaar toe met mogelijke schaalvoordelen als gevolg, maar gelijktijdig neemt de complexiteit veel meer toe dan bij horizontale integratie het geval zou zijn geweest, gezien de grote wettelijke en contextuele verschillen tussen de onderwijssectoren. De structuur van dergelijke besturen is weliswaar als divisioneel te beschouwen, met onderscheid naar de sectoren, maar het is de vraag of bundeling in een instelling niet een zodanige toename van de complexiteit tot gevolg heeft, dat het uiteindelijke effect negatief is.

Dit leidt tot de volgende vraag voor het onderzoek:

Creëert verticale integratie van schoolbesturen waarde?

Deze vraag is omgekeerd aan de vraag die in de vermelde literatuur aan de orde is, namelijk vast te stellen welke determinanten bepalend zijn voor welke vorm van governance. In dit onderzoek is de governance zelf gegeven en wordt nagegaan of dit een gunstige vorm is. Voor het onderzoek kan deze vraag vervolgens worden uitgesplitst in de onderstaande, te toetsen hypothesen.

Hypothesen:

1. Verbinding van vo met po leidt tot een gunstiger wervingspositie van leerlingen van geïntegreerde versus niet-geïntegreerde instellingen.
2. Verbinding van vo met mbo leidt tot een gunstiger wervingspositie van leerlingen van geïntegreerde versus niet-geïntegreerde instellingen.
3. Verbinding van po met vo leidt tot betere onderwijsresultaten voor leerlingen van geïntegreerde versus niet-geïntegreerde instellingen.

4. Verbinding van vo met mbo leidt tot betere onderwijsresultaten voor leerlingen van geïntegreerde versus niet-geïntegreerde instellingen.
5. Verticale geïntegreerde besturen hebben een gunstiger financiële positie dan niet-geïntegreerde instellingen.

4.6 Kwantitatieve analyse

4.6.1 Onderzoeksdata

Voor het beoordelen van de financiële prestaties van verticale instellingen is gebruik gemaakt worden van de financiële data van DUO (DUO, 2015). Deze bevatten jaarrekeninggegevens van de laatste vijf jaar.

Multisectorale instellingen zijn verplicht een deel van hun financiële gegevens te verbijzonderen naar de onderwijssectoren, de zogenaamde segmentatie. Speciaal gaat het daarbij om de rijksbijdragen. Uit de verdeling van de rijksbijdragen over de sectoren kan verticaliteit worden afgelezen en worden gekwantificeerd. Voorts is hierbij gebruik gemaakt van gegevens van de inspectie met betrekking tot de beoordeling van het financieel functioneren van instellingen.

Voor het analyseren van de doorstroomgegevens van leerlingen is gebruik gemaakt van de gegevens uit BRON (Basisregister Onderwijs), eveneens van DUO. In BRON worden leerlingen geregistreerd op basis van hun burgerservicenummer bij scholen (met een BRIN-nummer). DUO publiceert deze gegevens onder de OpenData voor respectievelijk po-vo, vmbo-mbo en havo/vwo-mbo (DUO, 2015). Daarmee kan inzichtelijk worden gemaakt of doorstroming binnen verticale besturen afwijkt van overige besturen.

Voor het analyseren van resultaatgegevens is gebruik gemaakt van de gegevens over het voortijdig schoolverlaten die eveneens via DUO beschikbaar zijn (DUO, 2015). Deze gegevens verschaffen inzicht in de mate waarin instellingen meer of minder succesvol zijn in hun doorstroom en kwalificatie van leerlingen.

Voor het analyseren van de personeelsgegevens is tenslotte ook gebruik gemaakt van de gegevens van DUO (2015). Het betreft hier gegevens over de samenstelling van het personeelsbestand over de periode 2009-2013 van alle besturen. In dat bestand is een onderscheid gemaakt tussen functiecategorieën directie, onderwijsondersteunend personeel en leraren (zowel regulier als in opleiding).

4.6.2 Beschrijvende statistiek

In het onderzoek zitten 1476 schoolbesturen. Daarvan zijn er 63 in enigerlei vorm verticaal. In tabel 4.2 staat de samenstelling ervan weergegeven. Omdat de categorieën elkaar deels overlappen, tellen de getallen niet rechtstreeks op tot de hierboven genoemde totalen.

Tabel 4.2: Besturen en het verticale karakter

	po	vo	mbo	hbo
N	1139	346	66	2 ²⁸
Verticaal po vo	36	36		
Verticaal po mbo	2		2	
Verticaal vo mbo		25	25	
Verticaal vo mbo hbo		2	2	2

Verticale besturen zijn, zoals te verwachten is, gemiddeld groter dan niet verticale. Verder zitten daar ook nogal wat verschillen tussen de sectoren in. Tabel 4.3 laat per onderscheiden categorie de gemiddelde balanstotalen zien, onderverdeeld naar besturen zonder en met een verticale structuur. De significantie van de verschillen is beoordeeld met een T-toets. Ook in deze tabel is sprake van overlap tussen de categorieën zodat de deeltotalen hoger zijn dan het totaal aantal besturen. Een verticaal bestuur po en vo komt zowel bij de besturen met po als bij de besturen met vo voor. Aangezien de combinatie met hbo slechts bij twee instellingen (gecombineerd met vo en mbo) voorkomt, is deze hier niet apart beschreven. Zoals de tabel laat zien, zijn verticale besturen in het funderend onderwijs (po en vo) aanzienlijk groter dan de overige. In het mbo maakt dat nauwelijks enig verschil.

Tabel 4.3: Grootte besturen in balanstotaal

	N	Gemiddeld balanstotaal niet verticaal (in miljoen)	Gemiddeld balanstotaal verticaal (in miljoen)	P<
Alle besturen	1476	7,4	42,7	0,000
Besturen met po	1139	3,8	22,1	0,000
Besturen met vo	346	12,7	42,4	0,000
Besturen met mbo	66	73,2	72,6	0,967

Verticale besturen komen verhoudingsgewijs het meest voor bij de Algemeen bijzondere schoolbesturen. Een vergelijking onder de vo-

²⁸ Dit zijn alleen de hbo-besturen in combinatie met een van de andere, onderzochte onderwijssoorten. De overige, alsmede combinaties met wetenschappelijk onderwijs, zijn buiten beschouwing gelaten.

besturen laat zien dat deze besturen daar maar 14 procent van de populatie vormen, terwijl zij ruim 40 procent van het verticale deel van de populatie zijn. Zoals eerder al is opgemerkt, omvat het Algemeen bijzonder onderwijs een zeer divers geheel van vaak onderwijskundige stromingen. Alle overige denominaties scoren lager dan hun proportie in de populatie, waarbij het protestants-christelijk onderwijs nog relatief het hoogst is met 27 procent van de populatie en 20 procent van het verticale deel.

4.6.3 Wijze van analyseren

De analyse van de wervingspositie (hypothesen 1 en 2) vindt plaats door vergelijking van de instroom van leerlingen in de eigen school met die van andere scholen. Daarbij wordt met behulp van een Herfindahl-Hirschman-index bekeken of de concentratie van aanleverende scholen bij verticale besturen significant verschilt van die bij monosectorale besturen. Een vergelijkbare analyse vindt plaats aan de uitstroomkant waarbij wordt gekeken of leerlingen vaker naar scholen van het eigen bestuur gaan. Bij beroepsonderwijs is het daarbij nog interessant te bezien in hoeverre de verticale doorstroming plaatsvindt in de beroepskolom. Hierbij wordt gekeken in hoeverre leerlingen vanuit een bepaalde stroom in het vmbo (bijvoorbeeld techniek) verticaal doorstromen naar de verwante opleidingen in het mbo.

De analyse van wervingsvoordelen in het beroepsonderwijs (hypothese 3) vindt plaats door enerzijds een vergelijkbare analyse te maken als in het primair en voortgezet onderwijs, waarbij wordt gekeken of de concentratie van toeleverende opleidingen afwijkt van die bij andere instellingen. Daarnaast is interessant te bezien hoe het zit met de verticale doorstroming langs de lijn van de beroepskolom. De analyse van de resultaatvoordelen ten behoeve van hypothesen 3 en 4 vindt plaats door vergelijking van het voortijdig schoolverlaten van instellingen met en zonder een verticale component.

De analyse van mogelijke financiële voor- of nadelen (hypothese 5) vindt plaats door vergelijking van de groep verticale (geïntegreerde) besturen met een groep waarbij dat niet het geval is. Hierbij is het zaak zorgvuldig de vergelijkbaarheid van de besturen in acht te nemen. Het zal, naast de verticale besturen, moeten gaan om een groep uit elk van beide sectoren waarover het multisectorale bestuur zich uitstrekt, en die op een aantal kenmerken met elkaar overeenstemmen (bijvoorbeeld grootte van het bestuur uitgedrukt in totale baten). De vraag bij deze analyse is of er significante verschillen zijn tussen de monosectorale en de multisectorale

besturen in het wel of niet voldoen aan een aantal financiële signaleringswaarden.

4.6.4 Doorstroom van leerlingen

Verticale integratie zou van invloed kunnen zijn op de doorstroming van leerlingen tussen de verschillende sectoren van het onderwijs. Bezien vanuit de positie van de aanleverende partij (de voorafgaande schakel in het sectorbouwwerk) zou dat betekenen dat scholen in een verticaal geïntegreerde vormgeving relatief meer leerlingen laten doorstromen naar de eigen vervolgscholen binnen het bestuurlijke verband. Bezien vanuit de positie van de ontvangende partij (de opvolgende schakel dus) betekent het dat scholen in een verticale context relatief meer leerlingen ontvangen van de eigen scholen dan van andere scholen.

Deze doorstroom is onderzocht door toepassing van de Herfindahl-Hirschman-index (HH-index). Deze index is een maat voor de marktconcentratie in een bedrijfstak (Calkins, 1983) en wordt ook gebruikt door de Nederlandse Zorgautoriteit (2015). De algemene vorm voor de berekening van de index is:

$$I = \sum_{i=1}^n M_i^2$$

Hierbij is M het relatieve marktaandeel. In een markt met een gegeven aantal deelnemers is de index gelijk aan de som van het kwadraat van de relatieve marktaandelen van die deelnemers. Toegepast op scholen betekent dit dat het relatieve belang van alle afnemende dan wel toeleverende scholen wordt gebruikt om een concentratiemaat voor de uitstroom respectievelijk instroom te berekenen. Daarbij is de veronderstelling, in aansluiting op de hypothesen, dat bij scholen in een geïntegreerd bestuurlijk verband een grotere concentratie te vinden zal zijn, namelijk tussen de scholen van het eigen bestuur. Zij zullen dus een hogere score op de index hebben.

De onderstaande analyses zijn gemaakt voor po-vo op basis van doorstroomgegevens over de jaren 2012 tot en met 2014 (DUO, 2015)²⁹. Daarvan waren 143.489 overstapwaarnemingen die betrekking hadden op 596.414 leerlingen beschikbaar (elke overstapwaarneming kan betrekking

²⁹ Het gaat per jaar steeds om leerlingen die een jaar eerder bij de voorafgaande sector ingeschreven stonden, bijvoorbeeld 2011 po – 2012 vo.

hebben op een of meer leerlingen). Voor vmbo-mbo zijn zij gebaseerd op de doorstroomgegevens van 2010 tot en met 2014 met 270.920 overstapwaarnemingen van in totaal 351.705 leerlingen (DUO, 2015). Deze gegevens zijn allereerst geaggregeerd van schoolniveau naar besturen. Waar sprake is van overstap tussen gelijke besturen, is sprake van verticaliteit, dus institutionele, verticale integratie. Uit de gegevens van de besturen is vervolgens de index berekend.

De analyse vanuit het po naar het vo is te zien in tabel 4.4. De besturen waar sprake is van integratie laten een lagere, gemiddelde indexwaarde zien dan de niet-geïntegreerde. PO-scholen die niet in een verticaal verband zitten, hebben dus een hogere concentratie van hun uitstroom. Dit is tegen de verwachting in. Het verschil tussen deze groepen is niet significant (t-toets, $p < 0,406$).

Tabel 4.4: HH-index voor uitstroom van po naar vo

Verticaal	N	Gemiddelde HH	Standaarddeviatie
Nee	932	0,4390	0,23050
Ja	31	0,4041	0,19754

Om wat meer inzicht te krijgen in deze resultaten, zijn de scholen met HH-indexwaarde 1 nader bekeken. Dat zijn er elf en daarvan zijn dus alle leerlingen in de onderzoeksperiode naar dezelfde school voor vo gegaan. Het overgrote deel daarvan kiest vanuit een sterk denominatieve insteek. Meer dan de helft bestaat namelijk uit een groep reformatische basisscholen in Zeeland waarvan alle kinderen naar dezelfde school voor vo gaan (die niet onder hetzelfde bestuur valt als de basisscholen). In enkele gevallen is er sprake van een regionaal monopolie van een vo-bestuur.

Bezien vanuit het vo, waar dus de concentratie van toeleverende po-scholen wordt beschouwd, is het beeld enigszins anders (tabel 4.5). Hier is de concentratie van toeleverende scholen bij verticale besturen wel groter dan bij de niet-verticale besturen. Opvallend is hierbij het grote verschil in variantie tussen de scores. Mede daardoor is het verschil van de gemiddelde scores niet significant (t-toets, $p < 0,098$).

Tabel 4.5: HH-index voor instroom in vo vanuit po

Verticaal	N	Gemiddelde HH	Standaarddeviatie
Nee	308	0,2162	0,12539
Ja	31	0,2925	0,24583

Een nadere beschouwing van de vier besturen met indexscore 1, waarbij dus alle leerlingen van basisscholen onder eenzelfde bestuur komen, laat

zien dat hier primair geografische omstandigheden een rol spelen. Twee van de vier zijn scholen op de Waddeneilanden en de derde op Urk, waar het niet voor de hand ligt dat kinderen van andere dan de plaatselijke basisscholen het voortgezet onderwijs komen bezoeken. Daarnaast scoort het joodse onderwijs hoog, waarbij de denominatieve richting dus een belangrijke rol speelt. Hypothese 1 wordt door deze analyses dus niet bevestigd.

Dezelfde analyses zijn uitgevoerd op de overstap van vmbo naar mbo. Hierbij is een correctie op de samenstelling van het scholenbestand aangebracht, aangezien niet alle vo-scholen vmbo aanbieden en het resultaat daardoor vertekend zou worden. Voor de analyse zijn daarom de scholen van het type SSG2 weggelaten³⁰. Tabel 4.6 laat de resultaten zien vanuit de vmbo-scholen. De niet-geïntegreerde besturen laten een hogere concentratie van hun uitstroom zien dan de geïntegreerde scholen. Het resultaat is niet significant (t-toets, $p < 0,268$).

Tabel 4.6: HH-index voor uitstroom van vmbo naar mbo

Verticaal	N	Gemiddelde HH	Standaarddeviatie
Nee	247	0,3905	0,16579
Ja	25	0,3517	0,17599

De concentratie laat bij deze overstap minder extreme waarden zien dan tussen po en vo. De belangrijkste overwegingen lijken hier geografische te zijn. De hoogst scorende vmbo-besturen liggen alle in de grensgebieden van het land, waar er eigenlijk maar een reële keuze is voor een instelling voor mbo om naartoe te gaan omdat aan de andere kant het buitenland of de zee ligt. Een regressieanalyse met de HH-index als afhankelijke variabele laat voorts zien dat het relatief lage resultaat bij de geïntegreerde besturen voor een belangrijk deel wordt verklaard door de doorstroming in de aoc's (agrarische opleidingscentra). Daarop zal nog nader ingegaan worden.

Bezien vanuit de ontvangstkant, het mbo, is het resultaat vergelijkbaar (tabel 4.7). Hier is er een iets hogere concentratie van de instroom bij geïntegreerde besturen maar ook hier is het verschil niet significant (t-toets, $p < 0,405$). De HH-scores zijn hier aanzienlijk lager dan in de andere gevallen (maximaal 0,55). Dat laat zien dat de mbo-instellingen (merendeels ROC's) hun leerlingen uit een groot voedingsgebied halen. De sterkste concentratie is een gevolg van een denominatieve component, namelijk een reformatorisch roc dat vooral van reformatorische scholen onder het eigen bestuur leerlingen ontvangt.

³⁰ Dit is de scholenindeling die door DUO wordt gebruikt. SSG2-scholen kennen alleen havo en vwo.

Tabel 4.7: HH-index voor uitstroom van vmbo naar mbo

Verticaal	N	Gemiddelde HH	Standaarddeviatie
Nee	41	0,1346	0,10846
Ja	25	0,1579	0,11190

Om nader zicht te krijgen op de verticale doorstroming in de beroepskolom, is een vergelijking gemaakt tussen typisch verticale instellingen en de rest. Typisch verticale instellingen zijn agrarische opleidingscentra (aoc's) en daarnaast enkele vakscholen, zoals het grafisch onderwijs. Kenmerk is dat zij in een specifieke beroepscategorie opereren. Tabel 4.8 laat zien dat de HH-index in het mbo betrekkelijk weinig verschilt tussen deze instellingen (t-toets, $p < 0,589$). Bij het vo is de concentratie juist tegengesteld aan wat verwacht zou mogen worden, al is het verschil net niet significant (t-toets, $p < 0,057$).

Tabel 4.8: HH-index voor in- en uitstroom naar typisch verticale instellingen

	Typisch verticaal	N	Gemiddelde HH	Standaarddeviatie
Vmbo naar mbo	Nee	258	0,3914	0,16527
	Ja	14	0,3043	0,17946
Mbo van-uit vmbo	Nee	52	0,1396	0,11373
	Ja	14	0,1576	0,09457

Om nog wat meer inzicht te krijgen in de verschillen, is een uitsplitsing gemaakt binnen de typisch verticale instellingen tussen aoc's en vakinstellingen. Tabel 4.9 laat het resultaat daarvan zien. Duidelijk wordt daaruit dat de lage concentratie duidelijk zit in de groep van aoc's. Zij vertonen een veel lagere concentratie dan bij de overige instellingen het geval is ($p < 0,000$). De vakinstellingen hebben juist een tegengesteld beeld van een veel hogere concentratie. Dat verschil is niet significant (t-toets, $p < 0,150$).

Tabel 4.9: HH-index voor uitstroom van aoc en vakinstellingen versus overig vmbo

	Verticaal	N	Gemiddelde HH	Standaard-deviatie
Aoc	Nee	263	0,3934	0,16549
	Ja	9	0,1993	0,06743
Vakinstelling	Nee	267	0,3850	0,16654
	Ja	5	0,4933	0,16163

De uitstroom vanuit het groene onderwijs is in het algemeen nogal afwijkend van die van de andere beroepssectoren (economie, techniek en zorg en welzijn). Tabel 4.10 laat de stroomgegevens tussen die richtingen zien van vmbo naar mbo. Merk daarbij op dat het enkel gaat om de lagere niveaus van het vmbo, waarin leerlingen al in een beroepsrichting zitten. De theoretische leerweg (mavo) is eruit weggelaten omdat die een algemeen karakter heeft en niet te koppelen is aan een beroepsrichting. Verder verschilt het aantal studenten tussen de verschillende richtingen van 28.405 (groen) tot 56.890 (zorg/welzijn) leerlingen. De tabel maakt duidelijk dat de doorstroom van groen naar groen veel lager is dan doorstroom in dezelfde richting in de andere sectoren. De HH-index die voor deze doorstroming is berekend, illustreert ook dat er van veel minder concentratie sprake is.

Tabel 4.10: proporties³¹ doorstroom van richtingen vmbo naar mbo

Van/naar mbo	Economie	Groen	Techniek	Zorg/welzijn	HH-index
Economie vmbo	0,67	0,02	0,12	0,15	0,48
Groen vmbo	0,24	0,31	0,18	0,26	0,25
Techniek vmbo	0,17	0,03	0,72	0,04	0,55
Zorg/welzijn vmbo	0,22	0,03	0,05	0,67	0,50

Tenslotte zijn twee analyses gemaakt van de instroom vanuit havo en vwo naar mbo. Dit is een minder gebruikelijke overgang die voor een deel gedaan wordt door ongediplomeerde uitstromers. In totaal gaat het om 35.566 leerlingen in de periode 2009-2014. Hierbij zijn van de vergelijking de scholen uitgesloten die geen havo/vwo hebben (SSG1). In tabel 4.11 staan respectievelijk de concentratie van de uitstroom vanuit havo/vwo en de instroom in het mbo. In beide gevallen zijn de verschillen niet significant (t-toets, $p < 0,821$ en $p < 0,630$).

Tabel 4.11: HH-index voor in- en uitstroom havo/vwo naar mbo

	Vert.	N	Gemiddelde HH	Standaarddeviatie
Havo/vwo naar mbo	Nee	258	0,3262	0,17895
	Ja	8	0,3407	0,12768
Mbo vanuit havo/vwo	Nee	58	0,1192	0,10359
	Ja	8	0,1378	0,08357

Dit alles levert dus geen bevestiging op van hypothese 2.

³¹ De proporties tellen niet allemaal op tot 1 in verband met kleine aantallen leerlingen die naar gemengde of onbekende richtingen zijn gegaan.

4.6.5 Voortijdig schoolverlaten

Om na te gaan of verticale integratie van invloed is op de opbrengsten van het leerproces zijn de gegevens over het voortijdig schoolverlaten (VSV) geanalyseerd. Daartoe is gebruik gemaakt van de gegevensbestanden omtrent voortijdig schoolverlaten in het vo en het mbo, beide over de periode 2005-2011 (DUO, 2015). Voor vo ging het daarbij om 6.391.096 ingeschreven leerlingen waarvan er 96.699 in die zeven jaar als VSV'er waren geregistreerd. Voor het mbo waren dat 2.706.589 studenten waarvan er 244.668 VSV'ers werden.

Voorafgaand aan de analyse is een controle uitgevoerd op de invloed van de denominatie openbaar. Zoals eerder al is aangegeven, zijn de meeste openbare instellingen in het eerste decennium van de eeuw overgegaan naar afzonderlijke rechtspersonen en is een deel daarbij in verticale besturen terechtgekomen. Het openbaar onderwijs kan leerlingen niet weigeren en dat zou van invloed kunnen zijn op de prestaties en dus het voortijdig schoolverlaten. In een regressievergelijking met VSV als afhankelijke variabele en verticaliteit en openbaar als verklarende variabelen, bleek dat verticaliteit wel ($p < 0,001$) en het openbare karakter geen ($p < 0,487$) significante invloed heeft op de prestaties.

Voor het vo zijn daarbij in principe twee groepen van verticale instellingen te onderscheiden, namelijk verticaal met primair onderwijs en verticaal met mbo. Allereerst is de analyse uitgevoerd op alle vo-instellingen met een verticaal karakter, ongeacht van welke aard. Dit levert de resultaten op in tabel 4.12. Het percentage VSV'ers is juist in de niet-verticale instellingen aanzienlijk lager dan in de verticale (t-toets, $p < 0,001$).

Tabel 4.12: VSV voor vo-besturen

Verticaal	N	Proportie VSV	Standaarddeviatie
Nee	256	0,0143	0,01958
Ja	49	0,0268	0,03516

Dit beeld verandert niet als de groepen worden gesplitst in combinaties met po respectievelijk mbo. In beide gevallen is het percentage VSV'ers veel kleiner bij besturen zonder verticale binding en zijn deze verschillen significant (t-toets, bij po: $p < 0,021$, bij mbo: $p < 0,018$) De hypothesen 3 en 4 worden hiermee dus niet bevestigd.

In het mbo is alleen een combinatie met vo geanalyseerd. Hier speelt het openbare karakter geen mogelijke rol van betekenis, zodat daarop niet afzonderlijk is gecontroleerd. In het mbo blijkt, zoals tabel 4.13 laat zien, het percentage VSV bij verticale instellingen iets lager te liggen dan bij de

andere. Het verschil is evenwel niet significant (t-toets, $p < 0,410$). De aantallen instellingen wijken in deze analyse iets af van de aantallen in de eerdere analyses. Dit wordt veroorzaakt doordat de gegevens lopen tot 2011. Nadien hebben zich enkele veranderingen in de samenstelling van het veld voorgedaan door de nasleep van Amarantis.

Tabel 4.13: VSV voor mbo-besturen

Verticaal	N	Proportie VSV	Standaarddeviatie
Nee	39	0,06072	0,00972
Ja	26	0,02579	0,00506

4.6.6 Financieel presteren

Verticale besturen zijn in het algemeen complexer dan niet-verticale. Dit kan onder meer van invloed zijn op de financiële prestaties. Daartegenover kan staan dat een dergelijk bestuur gemiddeld groter is, zodat van schaalvoordelen geprofitteerd zou kunnen worden. Om daarover een vergelijking te maken, is nagegaan welke besturen in de periode tot en met de jaarrekening van 2012 (geanalyseerd tot 1 juli 2014) onder aangepast financieel toezicht (AFT) van de inspectie hebben gestaan. Deze analyse is uitgevoerd over alle 1476 besturen. De onderstaande resultaten in tabel 4.14 laten zien dat er een groot verschil zit tussen beide. Van de verticaal geïntegreerde besturen heeft ruim 11 procent onder AFT gestaan, bij de niet-geïntegreerde is dat 2,5 procent. Dit verschil is significant (t-toets, $p < 0,037$).

Tabel 4.14: Besturen met aangepast financieel toezicht (AFT)

Verticaal	N	Proportie AFT	Standaarddeviatie
Nee	1413	0,0255	0,15763
Ja	63	0,1111	0,31769

Een deel van het negatieve financiële resultaat zou veroorzaakt kunnen worden door hogere kosten in de overhead. Dit is onderzocht met behulp van de personeelsbestanden van DUO (2015). Die bestanden hebben wel een aantal beperkingen. Allereerst bevatten ze alleen het personeel dat in eigen dienst van het bestuur is. Personeel in payroll-constructies of via uitzendbureaus valt daar dus niet onder. Een vergelijking van de ontwikkeling van het aantal fte's over 2009-2013 met de ontwikkeling van de personeelslasten (waar de lasten voor dergelijk personeel wel in terug moeten komen) laat zien dat in het po en vo die lasten per fte in vijf jaar met ruim 9 procent zijn gestegen. Aangezien er vrijwel die hele periode sprake was van een nullijn voor ambtenaren (waar ook

onderwijspersoneel onder viel) kan dat voor het overgrote deel alleen worden toegeschreven aan het inzetten van personeel in dergelijke constructies.

Omdat de gegevens over de personeelslasten echter geen informatie geven over de samenstelling van het personeelsbestand, is het toch noodzakelijk gebruik te maken van DUO-gegevens ondanks dat die niet op het hele personeelsbestand betrekking hebben. Daarnaast zitten er ook nog wel wat onnauwkeurigheden in de bestanden. Soms zitten er onverklaarbare breuken in de vijfjarige tijdreeksen en op sommige plaatsen is er sprake van onwaarschijnlijk hoge of lage overheadpercentages. Met inachtneming daarvan is in tabel 4.15 het verschil in percentage primair personeel (de tegenhanger van overhead) weergegeven. De tabel laat zien dat er een groot verschil is in het percentage overhead. Bij verticale besturen bedraagt dat ruim 31 procent. Bij de overige besturen is dat 24 procent. Dit verschil is significant (t-toets, $p < 0,000$)

Tabel 4.15: Proporties primair personeel

Verticaal	N	Primair personeel	Standaarddeviatie
Nee	1397	0,7589	0,09706
Ja	63	0,6843	0,09439

Onwaarschijnlijk hoge of lage percentages kunnen door fouten komen of door bijzondere omstandigheden. Als een school net in opbouw is, kan er korte tijd sprake zijn van 100 procent overhead (directeur die tegelijk de enige docent is). Dat kan echter niet alle uitschieters verklaren en het is niet mogelijk de cijfers inhoudelijk te controleren en corrigeren. Om die reden is de analyse nogmaals uitgevoerd met weglating van de bovenste en onderste 10 procent van de populatie. Dit levert dan, op grond van 80 procent van het totaal, het resultaat op in tabel 4.16. Het verschil is duidelijk een stuk kleiner, ruim 3,5 procent tussen verticale en niet-verticale besturen. Dat is nog altijd een significant verschil.

Tabel 4.16: Percentages primair personeel tussen percentiel 10 en 90

Verticaal	N	Primair personeel	Standaarddeviatie
Nee	1122	0,7685	0,05514
Ja	45	0,7330	0,05430

Het percentage overhead kan naast complexiteit door een verticale structuur ook afhankelijk zijn van de omvang van de instelling. Zoals eerder al duidelijk is gemaakt, zijn verticale instellingen in het algemeen groter dan de overige instellingen. De correlatie tussen verticaliteit en de natuurlijke logaritme van het balanstotaal bedraagt 0,245

(Spearman Rho, $p < 0,000$). Een regressievergelijking met het percentage primair personeel als afhankelijke variabele laat zien dat beide variabelen wel significant betekenis hebben voor de verklaring van het percentage overhead (verticaliteit $p < 0,000$, omvang $p < 0,006$). Zij verklaren echter slechts een zeer klein deel van de variantie ($R^2 = 0,021$).

Met dit alles is er evenwel geen bevestiging gevonden voor hypothese 5.

4.7 Conclusie en discussie

De intuïtieve keuze van de inspectie om verticaliteit of multisectoraliteit als risicofactor aan te merken, wordt door dit onderzoek in het algemeen bevestigd. De theorie rond verticale integratie laat zien dat er voor bekostigde schoolbesturen in Nederland eigenlijk geen goede redenen zijn om tot verticale integratie van onderwijssoorten over te gaan. Waar sprake is van een dergelijke verbinding, lijkt zij in de meeste gevallen min of meer toevallig tot stand te zijn gekomen. Marktoverwegingen spelen geen duidelijke rol en bedrijfseconomische overwegingen lijken vooral als gelegenheidsargument te worden gebruikt. Het resultaat onderbouwt de risicobenadering die de onderwijsinspectie in haar toezicht hanteert.

Deze mogelijk (bedrijfs)economische voordelen zijn nader in een empirische analyse onderzocht en blijken grotendeels niet te bestaan. Verticale besturen slagen er niet in leerlingstromen tussen de scholen van het eigen bestuur zo te laten verlopen dat er sprake is van een competitief voordeel. Schaalvoordelen die wel eens worden genoemd, komen in elk geval niet tot uitdrukking in het financieel functioneren van verticale instellingen. Zij doen het integendeel in dat opzicht juist veel slechter dan de andere besturen. De door de Onderwijsraad (2008) al geformuleerde vrees van een grotere bureaucratie lijkt in de omvang van de overhead inderdaad bevestigd te worden. De eventuele opbrengsten van een dergelijk samengaan zijn negatief, zodat er geen vraagstuk van winsttoewijzing maar juist van tekorten is. Er is al met al eigenlijk geen goede reden om verticale fusies in het Nederlandse onderwijs te laten plaatsvinden.

De vrees van Waterreus (2009) dat de verticale integratie leidt tot vermindering van de keuzevrijheid en sterkere doorstroming binnen het bestuur, wordt door de doorstroomanalyses niet bevestigd. Analyse van de in- en uitstroomgegevens van de besturen die wel respectievelijk niet geïntegreerd zijn over de sectoren, laat immers zien dat verticale verbinding niet leidt tot significant hogere doorstroming van leerlingen tussen scholen binnen het eigen bestuur. Als er sprake is van uitschieters,

zitten die, behalve in geografische beperkingen, vooral in een uitgesproken denominatieve richting en dan speelt dat veel meer een rol dan de vraag of de desbetreffende scholen zich wel of niet in een verticaal bestuurlijk verband bevinden. Ook tussen op zichzelf staande besturen doen zich die bewegingen in sterke mate voor.

Ook in het beroepsonderwijs heeft verticalisering geen duidelijke voordelen. De belangrijkste groep verticale instellingen, de aoc's, vertoont een significant lagere concentratie van doorstroming vanuit de (eigen) vmbo's naar het mbo dan bij de overige vmbo's het geval is. Dit probleem is overigens niet nieuw. Van den Berg (2013) heeft hier in opdracht van de aoc-raad recentelijk nog onderzoek naar gedaan. Hij stelt dat veel leerlingen vmbo's van aoc's kiezen omdat ze "klein, veilig, zorgzaam en wit" zijn en niet omdat hun belangstelling uitgaat naar een beroep in de groene sector. Die belangstelling weten de scholen in het verdere traject ook niet in voldoende mate te trekken. Alleen leerlingen die van huis uit een duidelijk beeld hebben van het groene beroep, kiezen hier bewust voor en zullen naar de verwachting van Van den Berg ook meer doorstromen naar het mbo. Dit is echter een kleine en nog altijd krimpende groep.

Waar het onderzoek geen antwoord op kan geven, is of de hybride vormen van samenwerking tussen in het bijzonder vmbo's en roc's winsten opleveren. Daarbij gaat het om onderdelen van instellingen die met elkaar vergeleken zouden moeten worden en waarbij heel specifieke stromen van leerlingen betrokken zijn. Met de beschikbare data is dat onmogelijk, maar het zou interessant zijn na te gaan of dergelijke trajecten leiden tot verbetering van de onderwijsopbrengsten en resultaten in de bedrijfsvoering. Een beperkt onderzoek hiernaar is uitgevoerd door Van Schoonhoven en Bouwmans (2013) in het kader van de evaluatie van de VM2-trajecten (zie paragraaf 5.4).

Literatuur

- Abu-Saad Huijjer, H. (1996). Verplegenden en verzorgenden; gekwalificeerd voor de toekomst. *Nederlands Tijdschrift Voor Geneeskunde*, 140(47), 2323-2325.
- Baker, G. P., Gibbons, R., & Murphy, K. J. (2002). Relational contracts and the theory of the firm. *Quarterly Journal of Economics*, 117(1), 39-84.
- Baker, G. P., Gibbons, R., & Murphy, K. J. (2008). Strategic alliances: Bridges between "islands of conscious power". *Journal of the Japanese and International Economies*, 22, 146-163.

- Barzel, Y. (1982). Measurement cost and the organization of markets. *Journal of Law and Economics*, 25(1), 27-48.
- Berg, N. v. d. (2013). *Doorstroom tussen groen vmbo en mbo. Onderzoek naar relevante factoren en sturingsmogelijkheden*. Rotterdam/Ede: Strix Aluco, AOC Raad.
- Calkins, S. (1983). The new merger guidelines and the Herfindahl-Hirschman index. *California Law Review*, 71(2), 402-429.
- Centraal Bureau voor de Statistiek (CBS). (2008). Verzelfstandiging openbare scholen zet flink door. Verkregen 08-21, 2015, van <http://www.cbs.nl/nl-NL/menu/themas/overheid-politiek/publicaties/artikelen/archief/2008/2008-90123-wk.htm>
- Commissie Kwalificatiestructuur Verpleging en Verzorging. (1996). *Gekwalificeerd voor de toekomst*. Zoetermeer/Rijswijk: Ministerie van Onderwijs, Cultuur en Wetenschappen/Ministerie van Volksgezondheid, Welzijn en Sport.
- Commissie onderzoek financiële problematiek Amarantis. (2012). *Autonomie verplicht. rapport onderzoek financiële problematiek Amarantis*. Den Haag:
- Coase, R. H. (1937). The nature of the firm. *Economica*, 4(16), 386-405.
- Commissie Fusietoets Onderwijs (CFTO). (2015). Adviezen scholenfusies in Nederland. Verkregen 07/10, 2015, van <http://cfto.nl/adviezen1>
- Dick, M. B., King, D. B., Mitchell, G. K., Kelly, G. D., Buckley, J. F., & Garside, S. J. (2007). Vertical integration in teaching and learning (VITAL): An approach to medical education in general practice. *The Medical Journal of Australia*, 187(2), 133-135.
- Dienst Uitvoering Onderwijs (DUO). (2015). Open databestanden onderwijs. Verkregen 07/11, 2015, van http://www.ib-groep.nl/organisatie/open_onderwijsdata/databestanden/default.asp
- Dienst Uitvoering Onderwijs (DUO). (2015a). Begrip: Denominatie. Verkregen 3 oktober, 2015, van https://duo.nl/includes/navigatie/openbare_informatie/waargebruikt.asp?item=Denominatie
- Geyskens, I., Steenkamp, J. E. M., & Kumar, N. (2006). Make, buy, or ally: A transaction cost theory meta-analysis. *Academy of Management Journal*, 49(8), 519-543.
- Grossman, S. J., & Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *The Journal of Political Economy*, 94(4), 691-719.
- Hendrikse, G. W. J. (2003). Transaction and influence costs. *Economics and management of organizations: Co-ordination, motivation and strategy* (pp. 203-240). London: McGraw-Hill.
- Hendrikse, G. W. J. (2004). Organisaties als systemen van attributen. *Tijdschrift Voor Economie En Management*, 49(3), 495-511.

Hendrikse, G. W. J., & Feng, L. (2012). Interfirm cooperatives. In A. Grandori (Ed.), *Handbook of economic organizations* (2013th ed., pp. 501-521) Edward Elgar.

Inspectie van het Onderwijs. (2015). *Handleiding primaire analyse en bijsluiters bij het bestuursoverzicht (GBT 1.1)*. Intern document.

Kooij, P. (2008). Het landbouwonderwijs in de twintigste eeuw. In M. Duijvendak, E. H. Karel & P. Kooij (Eds.), *Groen onderwijs. terugblik en uitzicht naar aanleiding van het 100-jarig bestaan van de vereniging voor hoger landbouw onderwijs (1906-2006)* (pp. 9-42). Groningen: Nederlands Agronomisch Historisch Instituut.

Langlois, R. N., & Robertson, P. L. (1989). Explaining vertical integration: Lessons from the American automobile industry. *Journal of Economic History*, 49(2), 361-375.

Mahoney, J. T. (1992). The choice of organizational form: Vertical financial ownership versus other methods of vertical integration. *Strategic Management Journal*, 13(8), 559-584.

Ménard, C. (2012). Hybrid modes of organization. alliances, joint ventures, networks, and other 'strange' animals. In R. Gibbons, & J. Roberts (Eds.), *The handbook of organizational economics* (pp. 1066-1108) Princeton University Press.

Ministerie van Onderwijs, Cultuur en Wetenschap (OCW). (2011). *Wijzigingswet diverse onderwijswetten (invoeren fusietoets in het onderwijs)*

Ministerie van Onderwijs, Cultuur en Wetenschap. (2014). In Ministerie van Onderwijs, Cultuur en Wetenschap (Ed.), *Kerncijfers 2009-2013*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.

Moerkamp, T., & Onstenk, J. (1999). In Cedefop, Europees Centrum voor de ontwikkeling van de beroepsopleiding (Ed.), *Beroepsonderwijs en scholing in Nederland* (2nd ed.). Thessaloniki: Cedefop, Europees Centrum voor de ontwikkeling van de beroepsopleiding.

Naaijkens, E. (2010, 24-09). Verticale fusie scholen in Waalwijk verkleint kloof tussen po en vo. Bericht geplaatst op <http://www.onderwijsbrabant.nl/content/verticale-fusie-scholen-waalwijk-verkleint-kloof-tussen-po-en-vo-0>

Nederlandse Zorgautoriteit. (2015). *Marktscan van de zorgverzekeringsmarkt 2015*. Utrecht: Nederlandse Zorgautoriteit.

Onderwijsraad. (2008). *De bestuurlijke ontwikkeling van het Nederlandse onderwijs*. (Nr. 20080309/930). Den Haag: Onderwijsraad.

Onderwijsraad. (2009). *Ontwikkelingsrichtingen voor het middelbaar beroepsonderwijs*. (Nr. 20090279/935). Den Haag: Onderwijsraad.

Pfeffer, J., & Salancik, G. R. (1978). *The external control of organizations: a resource dependence perspective*. New York: Harper & Row.

Powell, M. (2015). *An influence-cost model of organizational practices and firm boundaries*. Ongepubliceerd manuscript.

- Rosenthal, D., Worley, P. S., Mugford, B., & Stagg, P. (2004). Vertical integration of medical education: Riverland experience, South Australia. *The International Electronic Journal of Rural and Remote Health Research, Education Practice and Policy*, 4
- Schoonhoven, R. v., & Bouwmans, M. (2013). *Vijf jaar VM2. samenvatting van de uitkomsten van de VM2-monitor 2008-2012.* (No. 30082.07). 's-Hertogenbosch: ECBO.
- Tadelis, S., & Williamson, O. E. (2013). Transaction cost economics. In R. Gibbons, & J. Roberts (Eds.), *The handbook of organizational economics* (pp. 159-192). Princeton, New Jersey: Princeton University Press.
- Thomson, J. S., & Anderson, K. (2009). Reducing the load on GP teachers. *Australian Family Physician*, 30(11), 907-910.
- Waterreus, I. (2009). Is er een economische reden voor een fusietoets in het onderwijs? *TPEdigitaal*, 3(3), 80-102.
- Williamson, O. E. (1979). Transaction-cost economics: The governance of contractual relations . *Journal of Law and Economics*, 22(2), 233-261.
- Willik, John A. van der. (2010). Het laboratoriumonderwijs in Nederland: Verleden, heden en toekomst. *Analyse, Vakblad Biomedische Laboratoriummedewerkers*, 6, 171-174.

5 Conclusies en discussie

Dit hoofdstuk heeft een samengesteld karakter. Allereerst geeft het per onderzoeksonderwerp de centrale vraag en de conclusies die daaruit zijn getrokken. Het besteedt tevens aandacht aan onderzoek dat zou kunnen worden verricht om het onderwerp verder te verkennen. Vervolgens wordt in een aparte paragraaf aandacht besteed aan de vraag welke neveneffecten het gevolg kunnen zijn van de gehanteerde indicatoren. De laatste paragraaf heeft een op de praktijk gerichte inhoud. Tijdens het schrijven van dit proefschrift is de auteur actief betrokken geweest bij de ontwikkeling van het onderwijstoezicht. Dit heeft geleid tot directe, actieve toepassing van een aantal van de bevindingen. Deze worden hierin kort geschetst.

5.1 Beoordeling van de kwaliteitszorg

De hoofdstukken van dit proefschrift zijn gewijd aan het onderzoek van mogelijke indicatoren voor het risicogerichte toezicht door de onderwijsinspectie op onderwijsinstellingen. Deze indicatoren moeten beschouwd worden in het kader van een gelaagde principaal-agentrelatie tussen ministerie, toezichthouder en deze instellingen (Dixit, 2002).

Onderwijsinstellingen zijn, als agenten, verantwoordelijk voor de kwaliteit van hun onderwijs en worden geacht die kwaliteit systematisch te borgen. Daarbij hebben zij een grote mate van vrijheid om een en ander zelf in te richten. Zoals ook in hoofdstuk 1 al duidelijk is gemaakt, is de relatie met de principaal, het ministerie, er een van onvolledige contractering, die bovendien wettelijk is gewaarborgd door de grondwettelijke vrijheid, waaraan enkel bij wet beperkingen kunnen worden gesteld.

Ook de wijze waarop zij de kwaliteit borgen, is vrij. De onderwijswetgeving schrijft, zoals gesteld in hoofdstuk 2, geen modellen voor, maar spreekt in zeer algemene termen over de verantwoordelijkheid van besturen voor de borging van de onderwijskwaliteit. De Wet op het Voortgezet Onderwijs (art. 24 WVO) heeft het bijvoorbeeld over: "het beleid met betrekking tot de bewaking en verbetering van de kwaliteit van het onderwijs". De Wet op het primair onderwijs kent identieke bepalingen. In de Wet Educatie en Beroepsonderwijs (art. 1.3.6 WEB) wordt voorgescreven dat de instelling een stelsel van kwaliteitszorg inricht waarin voorzien wordt in een regelmatige beoordeling van de kwaliteit van het onderwijs. Het enige aspect daarvan (of misschien nog beter: voorwaarde daarvoor) dat zeer expliciet benoemd wordt, is de

bekwaamheid van het personeel. Het is opmerkelijk dat juist dat niet in de beoordeling van het kwaliteitszorgsysteem door de inspectie zit.

In hoofdstuk 2 is de vraag onderzocht in welke mate de gangbare beoordeling van kwaliteitszorg door de inspectie verklarend is voor het financieel en onderwijskundig functioneren van onderwijsinstellingen. Daarmee wordt duidelijk in hoeverre deze beoordeling kan dienen als een risicoindicator voor het toekomstig functioneren.

In het theoretische deel van het onderzoek is aangesloten bij het Baldrige-model (NIST, 1998). Dat is niet gebeurd omdat dat als een norm voor het kwaliteitszorgsysteem zou kunnen gelden, aangezien dat in strijd zou zijn met de vrijheid van de instellingen, maar omdat het een breed scala aan managementactiviteiten omvat en omdat er veel wetenschappelijk onderzoek is gedaan naar de effectiviteit ervan. Het empirische deel van het onderzoek is uitgevoerd met behulp van de standaardbeoordelingen die inspecteurs over een langere periode hebben uitgesproken over de kwaliteitszorg van de instellingen. Deze beoordelingen zijn gerelateerd aan de prestaties van de instellingen op het vlak van onderwijskwaliteit en financieel functioneren. Door de beoordelingen en de prestaties onder te verdelen in twee perioden, was het daarbij ook mogelijk na te gaan of er wellicht een oorzakelijke relatie was tussen de kwaliteitszorg in de eerste en de prestaties in de latere periode.

In het onderzoek is allereerst gebleken dat het “hebben” van een kwaliteitszorgsysteem op zichzelf niets zegt. Een dergelijke, algemene benadering kan niet gerelateerd worden aan de feitelijke kwaliteit. Vergelijking van de specifieke indicatoren van het inspectiekader met het kwaliteitszorgmodel van Baldrige laat zien dat dit kader slechts drie van de componenten bevat die volgens dat model relevant zijn en, zoals gezegd, juist niet het personeelsbeleid. Een positief beoordeelde kwaliteitszorg heeft zo geen uitstraling naar de organisatie waardoor deze in alle opzichten de kwaliteit borgt en verbetert.

De analyses in hoofdstuk 2 maken duidelijk dat die componenten hooguit voorspellende waarde kunnen hebben voor die onderwerpen in het functioneren van instellingen waarmee zij inhoudelijk direct verbonden zijn. Van een uitstralingseffect naar andere onderwerpen is niet merkbaar sprake. Onderwijskundige kwaliteitszorg zegt dus niets over het financieel functioneren en doet dat voor het onderwijskundig functioneren slechts voor zover het speciaal op aspecten van dat onderwijskundig functioneren is gericht, en dan alleen voor die aspecten. Dit sluit aan bij de opvattingen van Baker (2000) waar het gaat om uitlijning tussen doelen en de wijze waarop zij worden gemeten.

Dat heeft gevolgen voor het toezicht en het gebruik daarin van gegevens rond die kwaliteitszorg. Waar een kwaliteitsborgingsmechanisme in het kader van een risicogerichte systematiek informatief moet zijn over de onderwijskwaliteit, kan het dat alleen maar zijn als het veel specifieker wordt beschouwd en aldus een grotere uitlijning wordt gerealiseerd. In paragraaf 5.5 zal nader worden ingegaan op mogelijke gevolgen voor het toezicht.

De zeer uiteenlopende relaties tussen bestuurlijk opereren en de uiteindelijke prestaties, zowel onderwijskundig als financieel, met daartussen vele interferenties, bieden nog een groot onderzoeksterrein. Hooge (2013) maakt duidelijk dat dat geplaatst moet worden in de context van het omgaan met meerdere principalen en uiteenlopende doelen, zoals ook beschreven door Dixit (2002) en Shapiro (2005). Ook verder en breder onderzoek naar de feitelijke kwaliteitszorgsystemen in het onderwijs is wenselijk. Dat moet het inzicht vergroten in mate waarin en de uiteenlopende manieren waarop deze in het onderwijs worden ingericht, de doelen die ermee worden nagestreefd en speciaal de effecten die het oplevert. Daarbij is ook de vraag relevant in hoeverre de mogelijke toename van die systemen in de laatste jaren van invloed is geweest op de kennelijke toename van de gevoelde regeldruk bij het onderwijspersoneel.

Verder is er behoefte aan meer longitudinaal onderzoek naar de samenhang tussen financieel opereren door besturen en de gevolgen daarvan voor de onderwijskwaliteit. Onderzoek tot dusverre (Inspectie van het Onderwijs, 2014) laat zien dat financiële risico's niet een op een samenhangen met onderwijskwaliteit. De veronderstelling is ook hier dat het bestuurlijk opereren als achterliggende factor van relevante betekenis is en bovendien is er mogelijk sprake van een tijdfactor die cross-sectioneel onderzoek niet in beeld brengt. De Commissie onderzoek huisvesting ROC Leiden (2015) heeft laten zien dat in dit nogal extreme geval de onderwijskwaliteit na een aantal jaren de gevolgen ondervond van de financiële positie. Longitudinaal onderzoek naar de samenhang tussen financieel en onderwijskundig presteren zou daar meer inzicht in kunnen verschaffen.

5.2 Financieel disfunctioneren

Als principaal stelt de overheid aan de onderwijsinstellingen als agenten middelen ter beschikking om de na te streven doelstellingen te realiseren. Hieraan wordt de eis gesteld dat ze rechtmatig worden verworven en

rechtmatig en doelmatig worden besteed (OCW, 2011). Continuïteit van de publieke voorziening onderwijs is daarbij een belangrijke voorwaarde en die continuïteit komt af en toe ernstig in het geding. De toezichthouder wordt geacht daar, zo mogelijk preventief, namens de principaal toezicht op te houden.

De centrale vraag van het onderzoek in hoofdstuk 3 is daarom op welke wijze de toezichthouder continuïteitsrisico's bij onderwijsinstellingen eerder kan voorspellen om aldus tijdiger in te kunnen grijpen. Deze vraag is vanuit verschillende invalshoeken onderzocht.

Allereerst is onderzocht welke van vier verschillende modellen van financiële ratio's de beste resultaten gaven bij het voorspellen van risico's in het daarop volgende jaar. Vervolgens is nagegaan of een model, gebaseerd op gegevens van het direct voorafgaande jaar beter scoorde dan wanneer het gebaseerd was op een eerder jaar. Verder is onderzocht wat de optimale onderscheidingswaarde in de logistische regressie is, in hoeverre een hiërarchische beslissingsstructuur de voorspelling verbetert en of het daarbij uitmaakt of in die structuur gebruik wordt gemaakt van gelijksoortige of ongelijksoortige ratiomodellen. Dit alles is afgemeten aan de hand van het begrip accuraatheid, waarbij de verhouding tussen het aantal fouten van de eerste en tweede soort van belang is.

Als referentiemaat is daar niet, zoals bij soortgelijke onderzoeken, gebruik gemaakt van het al dan niet failliet gaan. Faillissementen komen, door de relatieve zekerheid van de bekostiging, in de bekostigde sector nog aanzienlijk minder voor dan in de marktsector die meestal wordt onderzocht. Bovendien is het beleid er actief op gericht die te voorkomen om de publieke dienstverlening niet in gevaar te brengen. Om die reden is als referentiepunt gekozen voor de formele toekenning van aangepast financieel toezicht aan de instellingen door de inspectie.

De analyses laten zien dat de verschillende ratiomodellen elkaar niet verontlopen in accuraatheid. De keuze voor een bepaald model is voor een goede detectie van risico's dus van minder belang. Wel van belang is het kiezen van een juiste onderscheidingswaarde in de regressie. Deze onderscheidingswaarde dient bij voorkeur overeen te komen met de proportie risicogevalen in de populatie. Een getrapte, hiërarchische besluitvorming met twee verschillende modellen verhoogt de accuraatheid van de voorspelling wel, maar met een hoge onderscheidingswaarde neemt het voordeel daarvan af. Er is geen voordeel aan het gecombineerd gebruik van modellen met een verschillend karakter, afkomstig uit de profit- en nonprofit-sector.

De voorspelling op basis van het laatste jaar scoort weliswaar beter dan die op basis van een jaar eerder, maar dat verschil is niet significant. Niettemin sluit deze waarneming wel aan bij de eerder genoemde waarnemingen van onder meer Rodriguez en Rodriguez (2006), Fuertes en Kalotychou (2006), Fantazzini en Figini (2009) en Duda en Schmidt (2010) waarin duidelijk wordt dat de financiële indicatoren in een laatste jaar voor een crisis dusdanige afwijkingen vertonen dat zij bruikbaar zijn voor korte-termijnvoorspellingen.

Het onderzoek laat zien dat het met een goede keuze van methoden mogelijk is tot 90 procent van de risicobesturen in een voorafgaand jaar correct te voorspellen. Daarmee is het zinvol dergelijke technieken, als onderdeel van een breder pakket in te voeren.

Omdat het onderzoek is ingericht op het op korte termijn voorspellen van de resultaten, is gekozen voor directe, logistische regressie. Het gebruik van discrete tijdanalyse wordt als methode geschikter geacht (Shumway, 2001) om door de analyses op de langere termijn meer diepgaand inzicht te krijgen in de mechanismen die ten grondslag liggen aan het financieel functioneren. Zeker voor een relatief jonge sector als het bekostigd onderwijs is verder onderzoek, gericht op het verkrijgen van dit inzicht, aan te bevelen.

Dat onderzoek zou dan enerzijds kunnen bijdragen aan verbetering van de beschreven detectiemodellen. Siau (2009) betoogde dat dergelijke modellen in het algemeen niet gebaseerd zijn op een theorie maar op empirische wijze via trial en error worden verkregen. Dat is ook in het onderzoek in hoofdstuk 3 het geval geweest. De benadering van Tuckman en Chang is daar in zekere mate een uitzondering op, waar het gaat om een bewuste keuze voor de toepassing van de inkomsten als een referentiemaat.

Buiten deze overwegend technische benadering op basis van kengetallen, zou onderzoek gericht op het ontwikkelen van een theorie over het financieel functioneren van bekostigde onderwijsbesturen in samenhang met de overige bedrijfsvoering en –inrichting wenselijk zijn. Garicano en Rayo (2016) laten voor een aantal grote bedrijven zien hoe keuzes die daarin worden gemaakt op het punt van de wijze van aansturing van personeel en de interne verdeling van autoriteit van doorslaggevende invloed kunnen zijn op het succes of het falen van die organisaties. Uiteindelijk is het succes op de langere termijn immers niet afhankelijk van de ontwikkeling van kengetallen, maar van de resultaten in het primaire proces. Om die reden zou in het verdere onderzoek de wisselwerking met de bestuurlijke aansturing van de organisatie en het

onderwijs, alsmede de onderwijskwaliteit in de beschouwing kunnen worden betrokken. Dit zou ook de inzichten in het bestuurlijk functioneren verdiepen.

5.3 Verticale integratie

In hoofdstuk 4 heeft een structuurkenmerk van onderwijsinstellingen centraal gestaan, namelijk de vraag of zij een verticaal verband vormen tussen onderwijsinstellingen. Dit sluit aan bij de constatering van de Commissie Amarantis (2012)³² waar zij de interim-voorzitter citeert die stelt: "het voortgezet onderwijs en mbo verschillen sterk van karakter; in het jarenlange streven naar doorlopende leerlijnen is dit ontkend, waardoor veel energie verkeerd gericht is geweest." Deze stelt tevens vast dat de fusies die daarvoor noodzakelijk waren, nooit echt waren gerealiseerd en dat onderlinge binding en gemeenschappelijk beleid zwak waren ontwikkeld.

Het onderzoek heeft zich daarom gericht op de beantwoording van de centrale vraag of verticale integratie van schoolbesturen waarde creëert. Deze vraag is onderzocht op een aantal terreinen, namelijk de doorstroming van leerlingen binnen de onderwijskolom, de uiteindelijke onderwijsresultaten van leerlingen en het financieel presteren van de instellingen. De conclusie daaruit is dat er eigenlijk geen argumenten zijn om verticaal te fuseren. Het levert geen betere wervingspositie op ten opzichte van andere schoolbesturen. De resultaten van de leerlingen in een verticaal geïntegreerd bestuur zijn niet beter dan in de andere besturen en het financieel presteren is duidelijk minder dan dat van de overige besturen.

Het onderzoek in hoofdstuk 4 heeft zich gericht op het verschil tussen de twee uitersten in het spectrum: volledige zelfstandigheid of institutionele integratie. Het hoofdstuk schetst echter tevens dat er nog verschillende andere vormen van governance kunnen zijn die daar tussenin liggen en die aan te duiden zijn als hybride vormen (Ménard, 2012). De aard en effectiviteit van dergelijke hybride vormen is, mede omdat de beschikbare data daartoe niet toereikend waren, buiten beschouwing gebleven.

Veel van de doorlopende trajecten van leerlingen zijn georganiseerd in dergelijke hybride vormen met relationele contracten. Afgezien van een recente evaluatie in het kader van de zogenaamde VM2-projecten (Schoonhoven en Bouwmans, 2013) en enige beschrijvende casuïstiek

³² Pagina 50.

rond doorlopende leerlijnen taal en rekenen tussen primair en voortgezet onderwijs (Münstermann en Hermans, 2010), is er nog weinig onderzoek gedaan naar de inrichting en de effectiviteit van andere vormen van bovensectorale aansturing van het onderwijs en van sectoroverstijgende leertrajecten. Deze beide studies gaan daar, gezien hun karakter, bovendien niet heel diep op in. Het zou daarom zeker de moeite waard zijn verder onderzoek te verrichten naar aansturingsvormen voor dergelijke trajecten.

5.4 Neveneffecten van de indicatoren

Sturing tussen principaal en agent, eventueel via een toezichthouder, kan gemakkelijk leiden tot neveneffecten, zoals in hoofdstuk 1 is betoogd. Naarmate de indicatoren daarvoor van groter belang zijn voor de beoordeling van de agent, zullen die verschijnselen sterker zijn. En naarmate de uitlijning tussen de door de principaal nagestreefde doelen en de wijze waarop zij gemeten worden, geringer is, zullen die effecten groter zijn (Baker, 2000). Voor de indicatoren die onderzocht zijn in de hiervoor beschreven onderzoeken, is dat niet anders.

Het “hebben” van een kwaliteitszorgsysteem als selectie-indicator door de toezichthouder kan tot dergelijke effecten leiden. Het onderzoek in hoofdstuk 2 heeft laten zien dat de uitlijning tussen de daar gekozen indicatoren en de werkelijkheid gering is. Als enkel het hebben (en eventueel toepassen) van een kwaliteitszorgsysteem van welke aard dan ook, een vrijwaring van risicogericht toezicht zou opleveren, is te verwachten dat dat zal leiden tot vormen van moral hazard (Dixit 2002) als window dressing, procedurele formalisering en teaching to the test. Bij het ontwerp van een toezichtstelsel dat daarop is gericht, zal daarom moeten worden gezocht naar een invulling die veel nauwer aansluit bij de doelstellingen van de principaal.

Bij de toepassing van risicoselectie op basis van een model met financiële indicatoren is dat eveneens het geval. Daarbij doet zich nog het bijzondere verschijnsel voor dat dergelijke indicatoren en de daarin gehanteerde signaalwaarden niet alleen fungeren als selectiemechanisme voor de toezichthouder. Doordat ze door de toezichthouder als agent van het ministerie worden gebruikt, krijgen ze ook een normatieve status voor de buitenwereld, onder meer voor financiers van de onderwijsinstellingen. De Commissie Amarantis (2012) heeft in haar rapport laten zien hoe de solvabiliteitsindicator van de inspectie daar als norm ging fungeren. Dit gebeurde zelfs zodanig dat de instelling kunstmatig bij de gehanteerde waarde wist aan te sluiten.

Dit verschijnsel kan beperkt worden door het kiezen van een groter aantal indicatoren die meer uiteenlopende aspecten in beeld brengen, maar daarbij zou een goede, theoretische fundering (Siau, 2002) niet mogen ontbreken. Bij ontbreken daarvan, bestaat het gevaar dat na incidenten een model ad hoc wordt uitgebreid om dat weer op te vangen.

De verticale integratie van een schoolbestuur heeft een heel andere werking als risicoindicator. Hier is sprake van een toestand of structuurkenmerk van het object van toezicht dat niet verbonden is met het handelen op korte termijn. Het is veeleer een contextvariabele die wel een indicatie is van een verhoogd risico, maar die op zichzelf ook geen aanleiding kan zijn voor actie door de toezichthouder. Dat zou wel het geval kunnen zijn in combinatie met andere risico's. Neveneffecten kunnen hierbij niet of veel minder aan de orde zijn.

5.5 Toepassingen in de praktijk

Sinds het onderzoek waarop hoofdstuk 2 is gebaseerd, hebben zich nieuwe ontwikkelingen voorgedaan in het toezicht op onderwijsinstellingen. Deze ontwikkelingen zijn beschreven in Inspectie van het Onderwijs (2015b). De inspectie beschrijft daarin dat zij in plaats van een vooral schoolgerichte onderzoeksbenadering gaat naar een bestuursgerichte onderzoeksbenadering. In deze benadering worden niet meer alle scholen onderzocht en worden zij in het algemeen niet onderzocht als op zichzelf staande entiteiten. De inspectie neemt zo meer afstand van het primaire proces zodat het risico optreedt van een geringere uitlijning (Baker, 2000). Bovendien wordt in zekere zin een extra principaal-agentrelatie tussengevoegd, namelijk die tussen bestuur en scholen.

De inspectie denkt daarin te voorzien door het onderzoek naar de sturing tussen bestuur en scholen een specifiek karakter te geven. Dit moet gebeuren in een voor de inspectie nieuwe onderzoeksvorm die erop gericht is samenhangende sturings- en controllijnen binnen het bestuur en zijn scholen op een aantal onderwerpen in beeld te brengen. Daarmee wordt feitelijk invulling gegeven aan het uitgangspunt dat kwaliteitszorg aangrijpingspunt is voor het toezicht op de kwaliteit van het onderwijs, maar ook aan het gegeven dat dit pas relevante informatie oplevert als het nauw aansluit bij die aspecten die daarin belangrijk zijn, zoals bevonden in hoofdstuk 2.

Om die reden vindt het onderzoek plaats aan de hand van een aantal thema's die van wezenlijk belang worden geacht voor de kwaliteit van het onderwijs en die daarom neergelegd zijn in kwaliteitsaspecten in het onderzoekskader van de inspectie. Bij elk van deze thema's onderzoekt zij het geheel van sturing, uitvoering en control op en tussen het centrale niveau van een instelling en de decentrale onderdelen. Daarmee wordt nauwer aangesloten bij de thema's die ertoe doen en vindt bovendien realiteitstoetsing plaats. Bovendien wordt, waar van toepassing, een relatie gelegd met de financiële aspecten van de sturing.

De realisatie van deze nieuwe vorm van toezicht zal vanaf 2017, als voorzien is dat deze wordt ingevoerd, aanzienlijke hoeveelheden gegevens op moeten leveren die kunnen bijdragen aan verder inzicht in het functioneren van besturen en het toezicht daarop. De inspectie verzamelt deze gegevens zelf, enerzijds om haar oordelen te onderbouwen en te publiceren en anderzijds in het kader van haar eigen kwaliteitszorg. In aansluiting op onderzoeken naar sturing in het onderwijs als dat van Hooge, Janssen, Look, Moolenaar en Sleegers (2015), zou dat een bron kunnen zijn voor verder onderzoek naar het functioneren van besturen en het gebruik van kwaliteitszorgsystemen daarbij.

In het financieel continuïteitstoezicht van de inspectie heeft toepassing van de analysetechnieken uit hoofdstuk 3 niet geleid tot invoering van een van de daar geschetste modellen. Daar gaven de bevindingen ook geen aanleiding toe. Wel was het aanleiding tot een kritisch onderzoek naar het praktisch nut van de aanwezige indicatoren in het risicomodel van de inspectie, zoals ook beschreven in hoofdstuk 3. Dit heeft geleid tot een verregaande bijstelling daarvan (Inspectie van het Onderwijs, 2015a). Met behulp van de logistische regressietechnieken die in hoofdstuk 3 zijn toegepast, zijn de bestaande risicoindicatoren afzonderlijk en in verschillende combinaties getoetst op hun relatie met de uiteindelijke vaststelling van aangepast financieel toezicht in verband met continuïteitsrisico's.

Uit deze analyse kwam, zoals ook hoofdstuk 3 al liet zien, naar voren dat vooral de solvabiliteit een sterk onderscheidende indicator is, en dat een deel van de gebruikte indicatoren als nutteloos was te beschouwen in de analyse. In sommige gevallen vertoonden de coëfficiënten een teken dat tegengesteld was aan de verwachtingen. Dit gold bijvoorbeeld voor de driejarige, negatieve rentabiliteit, waarvan intuïtief werd aangenomen dat deze een indicatie was voor financiële risico's omdat het bestuur dan krap zou komen zitten. Mogelijk speelt hierbij een rol dat non-profitinstellingen op langere termijn een rentabiliteit nul moeten hebben, en dat daar zowel positieve als negatieve rendementen deel van uitmaken.

Het onderzoek heeft geleid tot de opstelling van een nieuwe set indicatoren voor de detectie van risicovolle instellingen die leidt tot een hogere accuraatheid. Vrijwel gelijktijdig is een beleidsmaatregel genomen die instellingen verplicht in hun jaarverslag niet alleen de realisatiecijfers te geven van het verleden boekjaar, maar ook prognosecijfers voor de eerstvolgende drie jaar (OCW, 2007). Deze cijfers zijn zodanig van detail dat de gewenste indicatoren er ook voor de toekomst mee kunnen worden berekend en op deze wijze ook het toekomstig risico kan worden bepaald.

Het zou interessant zijn voor toekomstig onderzoek om na te gaan in hoeverre deze aanpassingen leiden tot een feitelijke verbetering van de detectie en tot vermindering van het aantal risicogeveallen. Ook is het nuttig te onderzoeken in hoeverre zich rond de prognosecijfers neveneffecten voordoen. Deze prognosecijfers vallen namelijk niet onder de controleverklaring van de accountant. Als zij aanleiding kunnen zijn tot intensievere aandacht door de toezichthouder, geeft dat ook aanleiding tot verschijnselen als window dressing. Vergelijking van de prognoses met latere realisatie kan daarin inzicht verschaffen en daarnaast ook in de bestuurlijke kwaliteit.

5.6 Algemeen

Toezicht in het algemeen en risicogericht toezicht in het bijzonder zijn, zoals in hoofdstuk 1 gesteld, afhankelijk van de juiste indicatoren voor de beoordeling van onder toezicht staanden of de detectie van risico's. Als deze indicatoren niet goed worden gekozen, is er onvoldoende sprake van uitlijning en doen zich legio kansen op ongewenste neveneffecten voor. De verschillende onderzoeken laten zien dat de keuzes voor dergelijke indicatoren om die reden een goede, wetenschappelijke fundering moeten hebben.

Intuïtieve aannames, zoals rond verticale instellingen, blijken soms te kloppen maar in andere gevallen volledig verkeerd te zijn, zoals het geval was met de indicatoren rond de meerjarige rentabiliteit. Zouden zich rond zo'n indicator effecten hebben voorgedaan, dan zouden die volledig tegengesteld zijn geweest aan de wenselijke uitkomsten. Een zorgvuldige verkenning van de mogelijkheden en risico's van verschillende indicatoren is dan ook wenselijk.

Daarbij ligt altijd het risico op de loer dat wapens worden gezocht om de vorige oorlog te winnen. Na de ondergang van ROC Leiden (Commissie onderzoek huisvesting ROC Leiden, 2015), ligt het zeer voor de hand op

zoek te gaan naar risicoindicatoren op het vlak van huisvesting. Zoals de Commissie echter heeft laten zien, was op het moment dat dergelijke indicatoren het risico zouden hebben gedetecteerd, reeds lang een punt gepasseerd waarop de toezichthouder met reguliere interventies het tij nog had kunnen keren. Om die reden is het wenselijk dat meer onderzoek wordt gedaan naar de onderliggende mechanismen dan direct gericht op momentane verschijnselen.

Literatuur

- Baker, G. P. (2000). The use of performance measures in incentive contracting. *The American Economic Review*, 90(2), 415-420.
- Commissie onderzoek financiële problematiek Amarantis. (2012). *Autonomie verplicht. rapport onderzoek financiële problematiek Amarantis*. Den Haag: Commissie onderzoek financiële problematiek Amarantis.
- Commissie onderzoek huisvesting ROC Leiden. (2015). *Ontspoorde ambitie*. Den Haag: Commissie onderzoek huisvesting ROC Leiden.
- Commissie Vermogensbeheer Onderwijsinstellingen. (2009). *Financieel beleid van onderwijsinstellingen. Rapport van de commissie vermogensbeheer onderwijsinstellingen*. Den Haag: Ministerie van OCW.
- Dixit, A. (2002). Incentives and organizations in the public sector: An interpretative review. *The Journal of Human Resources*, 37(4), 696-727.
- Duda, M., & Schmidt, H. (2010). *Bankruptcy prediction: Static logit model versus discrete hazard models incorporating macroeconomic dependencies*. (Unpublished Master). Lund University, Lund.
- Ehren, M. C. M., & Honingh, M. C. (2011). Risk-based school inspections in the netherlands: A critical reflection on intended effects and causal mechanisms. *Studies in Educational Evaluation*, 37, 239-248.
- Expertgroep Doorlopende Leerlijnen Taal en Rekenen. (2008). *Over de drempels met taal en rekenen*. Enschede: Stichting Leerplan Ontwikkeling (SLO).
- Fantazzini, D., & Figini, S. (2009). Random survival forests models for SME credit risk measurement. *Methodology and Computing in Applied Probability*, 11(1), 29-45.
- Fuertes, A., & Kalotychou, E. (2006). Early warning systems for sovereign debt crises: The role of heterogeneity. *Computational Statistics & Data Analysis*, 51, 1420-1441.
- Garicano, L., & Rayo, L. (2016). Why organizations fail: Models and cases. *Journal of Economic Literature*, 54(1), 137-192.
- Grossman, S. J., & Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *The Journal of Political Economy*, 94(4), 691-719.

Hooge, E. (2013). *Besturing van autonomie. over de mythe van bestuurbare onderwijsorganisaties*. Tilburg: Tilburg University.

Hooge, E., Janssen, S., Look, K. v., Moolenaar, N., & Slegers, P. (2015). *Bestuurlijk vermogen in het primair onderwijs. Mensen verbinden en inhoudelijk op een lijn krijgen om adequaat te sturen op onderwijskwaliteit*. Tilburg: TIAS School for Business and Society, Tilburg University.

Inspectie van het Onderwijs. (2014). *De staat van het onderwijs. onderwijsverslag 2012/2013*. Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs. (2015a). *Advies kengetallen*. Ongepubliceerd manuscript.

Inspectie van het Onderwijs. (2015b). *ONDERZOEKSKADER Ivho 2017. basisontwerp voor het toezicht op: Po, vo, (v)so en mbo*. Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs. (2015c). *De financiële situatie in het onderwijs*. Utrecht: Inspectie van het Onderwijs.

Ménard, C. (2012). Hybrid modes of organization. alliances, joint ventures, networks, and other 'strange' animals. In R. Gibbons, & J. Roberts (Eds.), *The handbook of organizational economics* (pp. 1066-1108) Princeton University Press.

Ministerie van Onderwijs, Cultuur en Wetenschap (OCW). (1963). *Wet op het voortgezet onderwijs (WVO)* [Wet van 14 februari 1963, tot regeling van het voortgezet onderwijs]

Ministerie van Onderwijs, Cultuur en Wetenschap (OCW). (1995). *Wet educatie en beroepsonderwijs* [Wet van 31 oktober 1995, houdende bepalingen met betrekking tot de educatie en het beroepsonderwijs]

Ministerie van Onderwijs, Cultuur en Wetenschap (OCW). (2007). *Regeling jaarverslaggeving onderwijs (RJO)* [Regeling van de Minister van Onderwijs, Cultuur en Wetenschap van 17 december 2007, nr. WJZ/2007/50507, houdende nadere voorschriften voor de inrichting van de jaarverslaggeving van door de Minister van Onderwijs, Cultuur en Wetenschap dan wel de Minister van Landbouw, Natuur en Voedselkwaliteit bekostigde onderwijsinstellingen.]

Münstermann, H., & Hermans, V. (2010). Doorgaande leerlijnen: Een complexe uitdaging. *Tijdschrift Onderwijsinnovatie*, 12(3), 35-37.

National Institute of Standards and Technology (NIST). (1998). *Malcolm Baldrige national quality award. 1998 criteria for performance excellence*. Milwaukee: American Society for Quality.

Rodriguez, A., & Rodriguez, P. N. (2006). Understanding and predicting sovereign debt rescheduling: A comparison of the areas under receiver operating characteristic curves. *Journal of Forecasting*, 25(7), 459-479.

Schoonhoven, R. v., & Bouwmans, M. (2013). *Vijf jaar VM2. samenvatting van de uitkomsten van de VM2-monitor 2008-2012*. (No. 30082.07). 's-Hertogenbosch: ECBO.

- Shapiro, S. P. (2005). Agency theory. *Annual Review of Sociology*, 31, 263-284.
- Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business*, 74(1), 101-124.
- Siau, C. (2009). De opstelling en het gebruik van een falingspredictiemodel. In D. Jaspers, & P. Vermoortel (Eds.), *Bedrijf & taal, opstellen voor Wilfried Janssens/Business & language, essays in the honour of Wilfried Janssens* (pp. 3-26). Leuven: Peeters.

Samenvatting

Nederlandse schoolbesturen zijn privaatrechtelijke entiteiten met overheidsbekostiging en een hoge mate van autonomie. De Inspectie van het Onderwijs houdt namens het ministerie van Onderwijs, Cultuur en Wetenschap toezicht op deze besturen en hun scholen. Dit toezicht heeft zowel betrekking op het onderwijskundig als op het financieel functioneren. Het toezicht kan beschouwd worden als een controlemechanisme in de principaal-agentrelatie tussen het ministerie en de schoolbesturen.

Het inspectietoezicht is risicogericht. Op basis van risicoanalyses van het onderwijskundig en financieel presteren door de besturen, kiest de inspectie systematisch haar onderzoeksmethoden en de middelen om deze uit te voeren. Daarbij maakt zij gebruik van risicoindicatoren. In dit proefschrift is een aantal van die indicatoren onderzocht op hun bruikbaarheid.

De eerst onderzochte indicator betrof de beoordeling door de inspectie van het kwaliteitszorgsysteem van de schoolbesturen. Besturen zijn verantwoordelijk voor de kwaliteit van hun onderwijs en zijn verplicht deze kwaliteit systematisch te borgen en verbeteren. Door het kwaliteitszorgsysteem van de besturen te beoordelen, hoopt de inspectie in staat te zijn genoeg informatie te krijgen over hun feitelijk presteren.

De onderzoeksvraag was daarom in welke mate de bestaande wijze van beoordeling door de inspectie van het kwaliteitszorgsysteem van de besturen hun financiële en onderwijskundige presteren kan verklaren. Het onderzoek heeft laten zien dat dat in beperkte mate het geval is. De beoordeling van de inspectie omvat veel relevante delen van een total-quality-management-systeem (het Baldrige Criteria for Performance Excellence model). Klant- en marktgerichtheid, procesmanagement en meting, analyse en kennismanagement zijn erin vertegenwoordigd. Maar aspecten als leiderschap, strategisch management en speciaal personeelsbeleid ontbreken in het model van de inspectie.

Het onderzoek laat zien dat alleen die aspecten in de beoordeling van het kwaliteitszorgsysteem die direct overeenkomen met overeenkomstige aspecten in de werkelijkheid, correlatie laten zien. Enkel het hebben van een kwaliteitszorgsysteem heeft geen relatie met het onderwijskundig of financieel presteren.

Het tweede onderzoek heeft zich gericht op de bruikbaarheid van financiële ratiomodellen voor het voorspellen van financieel (dis)functioneren door bekostigde instellingen. De onderzoeksvraag was of de inspectie in staat zou zijn deze eerder, efficiënter en effectiever te voorspellen. Er is een viertal financiële ratiomodellen met elkaar vergeleken. Twee van deze modellen (Altman en Ohlson) zijn ontworpen

voor bedrijven in de markt. Het derde model (Tuckman en Chang) was ontwikkeld voor non-profitorganisaties. Het vierde model was ontwikkeld door de inspectie zelf.

Deze modellen zijn met elkaar vergeleken onder verschillende omstandigheden om de optimale verhouding te bepalen tussen fouten van type 1 en type 2, uitgedrukt in het begrip accuraatheid. Dit hield de keuze in van de beste onderscheidingswaarde in de logistische regressies en het combineren van verschillende modellen in een soort hiërarchisch beslissingssysteem. De conclusie is dat verschillende modellen als zodanig elkaar niet ver ontlopen bij het voorspellen van de financiële kwetsbaarheid van schoolbesturen. Het combineren van verschillende modellen in een beslissingssysteem en de keuze van een onderscheidingswaarde die overeenkomt met de proportie financieel kwetsbare besturen in de populatie, kan leiden tot het voorspellen van 90 procent van de financiële problemen in een volgend jaar.

In het derde onderzoek is nagegaan in hoeverre verticale integratie van schoolbesturen meerwaarde heeft ten opzichte van besturen die zich enkel richten op een onderwijssector. In dit onderzoek is gekeken naar de wervingspositie van scholen, hun onderwijskundige resultaten en hun financiële prestaties. De conclusie hieruit is dat verticale integratie vanuit onderwijskundig oogpunt geen duidelijke meerwaarde heeft. In financieel opzicht vormt het een duidelijk risico. Verticale besturen hebben vaker financiële problemen en een grotere overhead.

Abstract

School boards of public education in the Netherlands are private entities with public funding and with a large degree of autonomy. Supervision of the boards is done by the Education Inspectorate of the Dutch Ministry of Education and involves both the educational and the financial performance of the boards. This supervision can be viewed as a control mechanism in the principal-agent relationship between the ministry and the school boards.

The supervision of the Inspectorate is risk-oriented. Based on risk assessments of the educational and financial performance of the boards, the Inspectorate systematically chooses its methods of supervision and the means to implement them. It does so by the use of risk indicators. In this thesis, the effectiveness of these indicators is examined.

The first indicator that is examined involves the assessment by the Inspectorate of the quality management system of school boards. School boards are responsible for the quality of the education they provide. They are obliged to systematically ensure and improve this quality. By assessing the quality assurance system used by the boards, the Inspectorate hopes to obtain sufficient information about their actual performance.

The research question is to which extent can the existing method of assessment by the Inspectorate of the quality management system of the school boards explain their financial and educational performance. The research shows that this is limited. The assessment by the Inspectorate captures several important parts of a total quality management system (the Baldrige Criteria for Performance Excellence Framework). Customer and market focus, process management and measurement, analysis and knowledge management are present. However, elements such as leadership, strategic planning and especially human resource management are missing in the model of the Inspectorate.

The research shows that only those elements in the assessment of the quality management system that match with the corresponding element in reality, are correlated with each other. Just having a quality system is not related to the educational or financial performance of the school boards.

The second study focuses on the usefulness of financial ratio models in predicting financial (dis)functioning by government funded educational institutions. The research question was whether the inspectorate would be able to predict these earlier, more efficiently and more effectively. Four

models of financial ratios were compared with each other. Two of these models (Altman and Ohlson) were developed for business companies. The third model (Tuckman and Change) was developed for non-profit organisations. The fourth model was developed by the Inspectorate itself.

These models are compared with each other under different conditions in order to determine the optimal ratio between errors of type 1 and type 2, expressed in terms of the accuracy. This includes the selection of the best cut value in the logistic regressions and combining the different models in a form of hierarchical decision making. The conclusion is that these models do not lead to very different results in predicting the financial vulnerability of school boards. Combining different models in a decision model and the selection of a cut value corresponding with the proportion of financially vulnerable boards in the population can lead to predicting up to 90 percent of the financial problems in a subsequent year.

The third study evaluates to which extent the vertical integration of school boards has added value compared to boards that only focus on education in one sector. The focus is on the recruitment position of schools, their educational outcomes and their financial performance. The conclusion is that vertical integration has no clear added value from an educational point of view. In financial terms, it is an obvious risk. Vertical boards often have financial difficulties and higher overhead.

Curriculum vitae

Jos Verkroost is geboren op 26 juni 1955 in Someren. Na een opleiding gymnasium-B aan het Bisschoppelijk College te Weert, behaalde hij in 1980 het doctoraal Nederlandse taal- en letterkunde aan de Katholieke Universiteit Nijmegen.

Vanaf 1978 was hij werkzaam als docent Nederlands in het mbo. Van 1983 tot 1996 heeft hij verschillende directiefuncties in het mbo vervuld.

In 1988 is hij gestart met een studie bedrijfskunde, later gecombineerd met bedrijfseconomie, aan de Open Universiteit. In 1994 heeft hij doctoraten behaald in bedrijfseconomie en bedrijfskunde.

Vanaf 1996 werkt hij als inspecteur en later als coördinerend inspecteur bij de inspectie van het onderwijs. Naast regulier toezicht heeft hij leidinggevende taken verricht. Sinds 2006 is hij voornamelijk betrokken bij innovatie van het toezicht en onderzoeken op toezichtgebied.

In 2013 is hij gestart als promovendus in een traject dat door de Erasmus Universiteit was geïnitieerd op het gebied van onderzoek naar toezicht. In het kader daarvan heeft zijn promotieonderzoek gedaan bij George Hendrikse.

Jos Verkroost is gehuwd en heeft twee kinderen.

Portfolio

- Broek, E. J. J. v. d., Kokmeyer, E., Koopmans, T., & Verkroost, J. J. H. (1985). *Nederlands voor het economisch onderwijs 1*. Culemborg: Educaboek.
- Broek, E. J. J. v. d., Kokmeyer, E., Koopmans, T., & Verkroost, J. J. H. (1986). *Nederlands voor het economisch onderwijs 2*. Culemborg: Educaboek.
- Broek, E. J. J. v. d., Kokmeyer, E., Koopmans, T., & Verkroost, J. J. H. (1987). *Nederlands voor het economisch onderwijs 3*. Culemborg: Educaboek.
- Hardeveld, G. v., & Verkroost, J. J. H. (2013). Een kwaliteitskader voor de onderwijsinspectie. *Kwaliteit in Bedrijf*, 4, 16-19.
- Hardeveld, G. v., & Verkroost, J. J. H. (2013). Kwaliteitssysteem bij de onderwijsinspectie. *Sigma*, 1, 12-15.
- Hardeveld, G. v., & Verkroost, J. J. H. (2013). Wat doet de inspectie zelf aan kwaliteitszorg? *Basisschool Management*, 27, 26-30.
- Janssens, F. J. G., & Verkroost, J. J. H. (2003). Proportioneel toezicht in een kennisgeoriënteerde inspectie. In B. Snel, & R. v. Naerssen (Eds.), *Toezicht gebundeld. actuele ontwikkelingen in het toezicht op het onderwijs* (). Utrecht: Inspectie van het Onderwijs.
- Roeters, A. S., & Verkroost, J. J. H. (2013). Toezicht op publieke belangen, een reactie. *Tijdschrift Voor Toezicht*, 4(4), 69-74.
- Smeets, G., & Verkroost, J. J. H. (2011). *Selectief en slagvaardig. werken met de WOT (2000-2010)*. Utrecht: Inspectie van het Onderwijs.
- Verkroost, J. J. H. (1984). In Koster F. S. (Ed.), *Examenbundel MEAO 1979/1984*. Culemborg: Educaboek.
- Verkroost, J. J. H., & Cuypers, L. H. F. (1994). *BV in balans. opleidingscertificaat ABV en uitstroomcertificaat FA-IB*. Voorschoten: Educatief B.V.
- Verkroost, J. J. H., & Jurna, S. (2001). *Self-regulation and quality assurance in dutch secondary vocational education*. (Contribution to the European Forum on Quality in VET.).Den Haag.
- Wolf, I. F. de, & Verkroost, J. J. H. (2011). Evaluatie van de theorie en praktijk van het nieuwe onderwijstoezicht. *Tijdschrift Voor Toezicht*, 2(2), 7-24.
- Verkroost, J. J. H., De Ruijter-Mooren, E., & Zeeman, A. (2014). Samenwerking rijk-gemeente bij het toezicht op de leerplichtwet. *Tijdschrift Voor Toezicht*, 5(1), 45-53.

Colofon

Realisatie: Uitgeverij BOXPress || Proefschriftmaken.nl

ISBN/EAN 9789462953765

